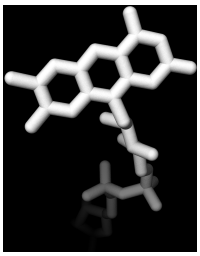
Perspectives in Yellow

Studies on Flavoprotein Monooxygenases

Anette Riebel

Cover design © Anette Riebel & Sebastian Bartsch, Leipzig (Germany) 2013



Front: modified picture from the installation *untitled* of **Dan Flavin** (1933 – 1996, American Minimal artist, creating light installations with off-the-shelf fluorescent tubes), taken 2012 in the Hirshhorn Museum in Washington, D.C.

Back: Autostereogram of FAD behind a pattern derived from the crystal structure of PAMO. Turn the book horizontal (solution see picture on the left).

The research described in this thesis was carried out at the Groningen Biomolecular Sciences and Biotechnology Institue (GBB) in the molecular enzymology group of the University of Groningen according to the requirements of the Graduate School of Science (Faculty of Mathematics and Natural Sciences, University of Groningen) and was supported by the Netherlands Organization for Scientific Research (NWO). It was part of the BIOMOX cluster of the "Integrated Biosynthesis and Organic Synthesis (IBOS)" programme of the NWO.

rijksuniversiteit groningen

PERSPECTIVES IN YELLOW

STUDIES ON FLAVOPROTEIN MONOOXYGENASES

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 22 november 2013
om 16.15 uur

door

Anette Riebel

geboren op 15 oktober 1983
te Worms, Duitsland

Promotor: Prof. dr. ir. M. W. Fraaije

Beoordelingscommissie: Prof. dr. W. J. H. van Berkel
Prof. dr. L. Dijkhuizen
Prof. dr. G. J. W. Euverink

ISBN (printed version): 978-90-367-6536-7
ISBN (electronic version): 978-90-367-6535-0

I wonder how, I wonder why
Yesterday you told me 'bout the blue blue sky
And all that I can see is just a yellow lemon-tree

Fool's Garden 1995, Lemon Tree

TABLE OF CONTENTS

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

Abbreviations

MO	monooxygenase
BVMO	Baeyer-Villiger monooxygenase
FMO	flavin-containing monooxygenase
NMO	N-hydroxylating monooxygenase
FAD	flavin adenine dinucleotide
PTDH	phosphite dehydrogenase
NAD(H)	nicotinamide adenine dinucleotide
NADP(H)	nicotinamide adenine dinucleotide phosphate
CRE	coenzyme regenerating enzyme
k_{cat}	catalytic constant
K_M	Michaelis constant
CE	cell extract
CCE	cleared cell extract
OYE	old yellow enzyme
UV	ultra violet
nm	nanometer
DMSO	dimethyl sulfoxide
MES	2-(N-morpholino)ethanesulfonic acid
aa	amino acid
e.g.	*latin: exemplī grātiā* (“for example”)
i.e.	*latin: id est* (“that is”)
et al.	*latin: et alii (m)/aliae (f)* (“and others”)
IBOS	Integrated Biosynthesis and Organic Synthesis
NWO	Netherlands Organization for Scientific Research

Enzymes are proteins, things of beauty and joy forever.

R.N. Perham 1976

CHAPTER 1

GENERAL INTRODUCTION:

FLAVOPROTEIN MONOOXYGENASES

Abstract: This chapter gives an overview of oxidative enzymes in general and the classification and characteristics of the flavoprotein monooxygenases in particular. The typical cofactors for flavoprotein monooxygenases, FAD and NAD(P)H, are described in detail. This chapter also provides information on the goal of the described research and the specific organism *Rhodococcus jostii* RHA1 which has been used as source of the enzymes described in this thesis.

1.1. Oxidative enzymes

Enzymes are well known and thoroughly studied biocatalysts that allow green and alternative solutions for the chemical industry. They are categorized in six main classes, depending on their chemical reactivity. Enzymes studied in this thesis belong to the class of oxidoreductases or redox enzymes (EC 1.x.x.x) (Dixon & Webb 1979). Such enzymes catalyze the transfer of electrons in biochemical reactions. Being more precise, the enzymes described here belong to the subclass of monooxygenases (EC 1.13.x.x and EC 1.14.x.x). These enzymes perform oxygenation reactions using a suitable electron donor and molecular oxygen (O_2) as oxygen donor. Hereby, monooxygenases insert only one single oxygen atom into the organic substrate under the elimination of water.

To perform the spin-forbidden reaction between molecular oxygen and an organic compound, the ground-state dioxygen in its triplet state needs to be activated first by receiving electrons. Therefore, most oxidative enzymes require cofactors like flavins, hemes or transition metal-ions for the catalytic activity (Cirino & Arnold 2002). The affinity of the cofactor towards the enzyme differs between:

- coenzymes that act like a substrate and interact with different enzymes, or
- a tightly (and sometimes even covalently) bound prosthetic group, which is part of the enzyme structure itself (Torres Pazmiño *et al.* 2010).

Regarding their chemical composition, cofactors can be divided in two main classes: inorganic (*i.e.* metal ions and iron-sulfur clusters) (Andreini *et al.* 2008), and organic cofactors (*i.e.* flavin and heme) (Fischer *et al.* 2010). In processes involving electron transfers mainly organic cofactors are used as 80% of all oxidoreductases employ such a cofactor (Kopacz *et al.* 2012).

1.2. The yellow cofactor: flavin

In 1932 the first evidence for the oxidative power of a yellow compound was found in brewers' bottom yeast by Warburg & Christian (1932) when they observed the biological oxidation of glucose-6-phosphate. In 1935 Theorell was able to purify the oxidative

compound, yielding a colorless apo protein and a yellow dye. Both components were shown to be necessary for the observed activity. This, for the first time indicated the essential role of proteins and their cofactors for biocatalysis. After a "new" yellow protein was found in yeast 1938 by Haas, the "old" one got its final name which is still valid today: old yellow enzyme (OYE) (Williams & Bruce 2002). Soon after its initial discovery it was realized that the yellow dye in OYE was similar to vitamin B_2 (riboflavin). This showed that vitamins have biochemical roles in enzymes. Theorell could show that the yellow cofactor in OYE was a phosphorylated form of riboflavin: riboflavin 5'-phosphate or flavin mononucleotide (FMN). He could also show that FMN could be bound again by the apo protein after purification, restoring the activity of the inactive apo protein (Theorell 1935, Ghisla & Edmondson 2009). For the discovery of the interaction between enzymes and their cofactors he was awarded the Nobel Price in Medicine in 1955. In the "new yellow enzyme", flavin adenine dinucleotide (FAD) was found to be the flavin cofactor and its chemical synthesis was honored with the Noble Prize in Chemistry to Sir Todd in 1957.

Together, FAD and FMN represent the two most common flavin cofactors containing the 7,8-dimethyl-10-alkylisoalloxazine ring system (Warburg & Christian 1933). Their biosynthetic precursor is riboflavin (vitamin B_2) (Kopacz *et al.* 2012) and FAD/FMN represent its biologically active form (Figure 1-1). In contrast to the aqueous solubility of their precursor riboflavin ($< 10^{-4}$ mol L^{-1}), the solubility of FMN and FAD is high (approx. 0.1 mol L^{-1}) due to the negatively charged phosphorylated side chain (Ghisla & Edmondson 2009).

The isoalloxazine ring system is a formal fusion of xylene (dimethylbenzene, A-ring), pyrazine (phenylenediamine, B-ring) and pyrimidine (C-ring) (Ghisla & Edmondson 2009) and is responsible for the amphipathic character. Mainly the xylene ring has been observed to be involved in substitution and addition reactions that result in naturally occurring flavin cofactor derivatives. Being the redox-active center of flavin cofactors, electrons are delivered to and removed from the three ring system, mainly via the central N_5 or bridgehead $C_{4\alpha}$ atoms (Walsh & Wencewicz 2013). In this way, reversible one- or two-electron transfer reactions, as well as dioxygen activation are possible (Massey 1994).

These functionalities, together with the fact that these reactions are kinetically and thermodynamically favorable, makes the flavin cofactor a powerful biological redox catalyst (Ghisla & Edmondson 2009, Walsh & Wencewicz 2013).

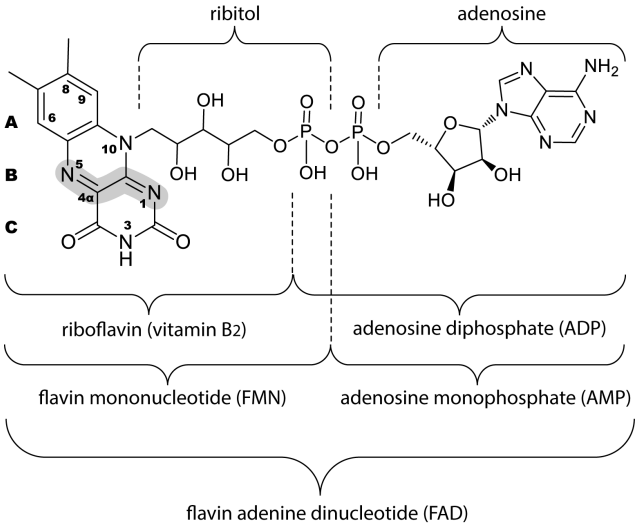


Figure 1-1: Structural formulas of riboflavin (B_2) and the flavin cofactors FMN and FAD.The electrophilic double imine moiety N_5-$C4_\alpha$-$C1_\alpha$-N_1 is highlighted in grey.

The central pyrazine is surrounded by a xylene and a pyrimidine. The pyrophosphate moiety and the adenine ring are used for binding and positioning of the flavin within the enzyme (Fruk *et al*. 2009, Kopacz *et al*. 2012). Also the hydrophilic pyrimidine ring allows interactions with the protein environment via hydrogen-bonding, acting as both, donor and acceptor. Such modulations of the pyrimidine ring can also influence the properties of the pyrazine ring. The hydrophobic xylene ring on the other side allows stacking with aromatic amino acid side chains in proteins (Ghisla & Edmondson 2009).

While the pyrimidine ring is relatively electron deficient, the xylene is more electron rich. The pyrazine ring is bridging these two systems and therefore its N_5 and $C_{4\alpha}$ show electrophilic properties, allowing photoaddition as well as nucleophilic addition reactions. Modifications in the pyrimidine or the benzene moiety can strongly affect the redox potential of the flavin and several natural occurring variants are known (*e.g.* 8-desmethyl *O*- or *N*-substituted flavin; 6- substituted flavin) (Ghisla & Edmondson 2009).

It is noteworthy that both names 'FMN' and 'FAD' are incorrect. FMN is not a mononucleotide, nor is FAD a dinucleotide since the bond between the ribose moiety and the isoalloxazine ring is not glycosidic. Nevertheless, both names are generally accepted in the scientific community and also by the IUPAC.

1.2.1. Yellow by sight

A prominent feature of the flavin cofactor is its bright yellow color that is also responsible for its name: *flavus* (latin = yellow). The yellow color results from the lowest transition state of the isoalloxazine chromophore around 450 nm. In total four prominent electronic transitions can be detected in the visible/ultraviolet (UV) spectral region (Figure 1-2).

The reduction of the flavin by one electron generates the flavin radical, the semiquinone form. In its neutral form it displays a blue color due to the absorption maxima at 615 nm and 580 nm, while in anionic form the color shifts to red (Massey & Palmer 1966). Adding another electron yields the 'leuco' or fully reduced flavin, which is colorless.

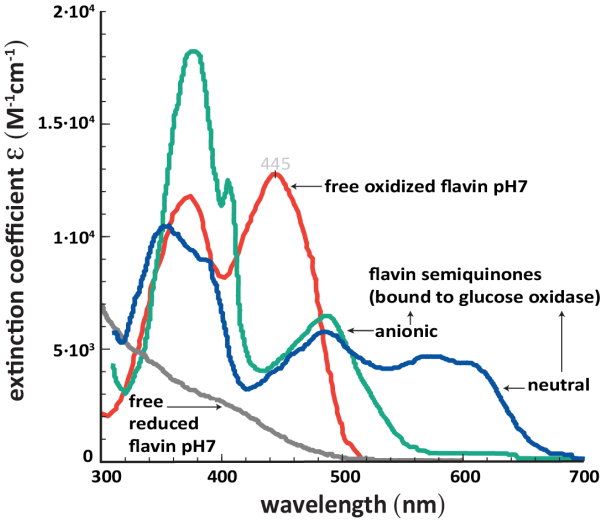


Figure 1-2: Absorption spectra of the flavin molecule in different redox states and ionic forms. The radical forms (semiquinones) can only be stabilized when bound to a protein. Spectra taken from Ghisla & Edmondson 2009 and modified.

As result of the yellow color of oxidized flavin, flavoproteins are typically yellow when isolated. This feature gives clear indications about the status of the expressed protein: a properly folded and flavin equipped holo flavoenzyme can be easily detected by sight. The flavin is not only absorbing light at specific wavelengths, it also displays typical fluorescence emissions. Therefore changes in redox and protonation states can also be observed in fluorescence emission spectra. Oxidized flavin emits light at 520-540 nm, while semiquinones (anionic or neutral) and the fully reduced flavins are colorless in aqueous solutions and poorly fluorescent. Differences in flavin fluorescence quenching are also used to distinguish between FMN and FAD which is caused by an intramolecular interaction between the isoalloxazine ring and the adenine ring moieties in FAD.

Both, the absorption spectra and the fluorescence emission of the different redox states and protonation levels of the flavin are strongly dependent on the environment the flavin is situated in. This allows probing the microenvironment around the flavin like ligands, protein, lipid membrane or solvent.

The obvious yellow color of a flavoprotein and the fact that the spectral properties of the flavin change with its redox state brought us to the idea to develop a method to determine the concentration of active enzyme in cell extracts: only enzymes with a properly incorporated flavin will be active and therefore can selectively be reduced by a suitable substrate. As proper flavin incorporation requires a properly folded enzyme environment, we can quantify the produced flavoprotein monooxygenases by using NADPH. The method and its outcome can be found in ***chapter 3****.*

1.2.2. Mechanism

In flavoprotein monooxygenases the catalytic cycle (Figure 1-3) can be divided in two half reactions. In the *reductive half reaction* (Figure 1-3, 5 → 1) the oxidized flavin becomes reduced by the oxidation of a cosubstrate (NAD(P)H). In the *oxidative half reaction*

oxygenation of the substrate takes places and the oxidized flavin cofactor is formed again. In this way the flavin has no impact on the reaction stoichiometry what makes it a real catalyst (Walsh 1980).

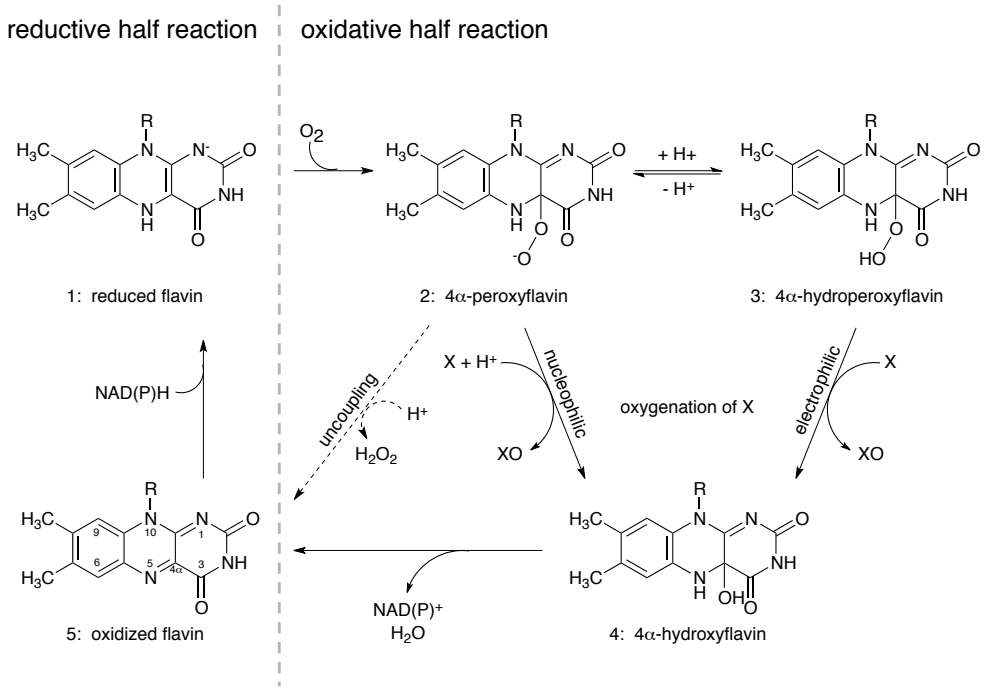


Figure 1-3: Catalytic cycle of flavin-dependent monooxygenases (van Berkel et al 2006).

At the beginning of the flavin-cycle the flavin gets reduced (5 → 1) by NAD(P)H via a hydride transfer to N_5 due to the electrophilic double imine moiety (*i.e.* N_5-$C_{4\alpha}$-$C_{1\alpha}$-N_1 in the oxidized flavin). Subsequently, the reduced and therefore electron rich flavin performs a one-electron transfer to molecular oxygen, creating a reactive intermediate (consisting of a superoxide and a flavin radical). A spin-inversion then allows the formation of a covalent bond between the oxygen and the $C_{4\alpha}$ of the isoalloxazine ring system yielding the $C_{4\alpha}$-peroxyflavin (1 → 2) (Ghisla & Massey 1989). This intermediate stays in equilibrium with its protonated equivalent, the $C_{4\alpha}$-hydroperoxyflavin (3) (Walsh & Chen 1988, Sheng *et al.* 2001). Even though both forms are stabilized within the enzyme, they will decay to the oxidized flavin (5) (Entsch & van Berkel 1995) in two possible ways:

either directly in the absence of a substrate under the formation of hydrogen peroxide (H_2O_2) as a net oxidase reaction (uncoupling, 2 → 5). Alternatively, an oxygenation will occur. The oxygenation of a substrate (2/3 → 4) results in formation of a $C_{4\alpha}$–hydroxyflavin (4). Such hydroxyflavin decays by release of a water molecule (4 → 5) (Massey 1994).

For the oxygenation reaction, both a nucleophilic or an electrophilic attack are possible, depending on the protonation state of the $C_{4\alpha}$-flavin intermediate. Which of the two peroxy-forms of the flavin (peroxyflavin (2) or hydroperoxyflavin (3)) is involved in catalysis strongly depends on the enzyme and/or substrate. A nucleophilic attack is required for Baeyer-Villiger reactions and epoxidation reactions (2 → 4), while an electrophilic attack is proposed for sulfoxidations and hydroxylations (3 → 4) (Walsh & Chen 1988, Chen *et al.* 1999, Colonna *et al*. 2003).

1.3. Priming the flavin cofactor: NAD(P)H as reducing coenzyme

To be able to perform oxygenating reactions, the flavin cofactor needs to be charged with electrons. Monooxygenases that are able to obtain these electrons from the substrate itself are called internal monooxygenases (EC 1.13.12) and are extremely rare. The majority of monooxygenases belong to the external family (EC 1.14.13) and depend on external electron donors like reduced coenzymes (van Berkel *et al.* 2006). The enzymes discussed in this thesis are all external monooxygenases and, as seen in Figure 1-3, specifically employ nicotinamide coenzymes.

Nature has developed two quite similar forms of nicotinamide cofactors (Figure 1-4), differing only in one phosphate group at the ribose 2´-position: NADH (nicotinamide adenine dinucleotide) and NADPH (nicotinamide adenine dinucleotide phosphate).

In 1906 the biochemists Harden and Young first discovered a substance (later identified as NADH) in “yeast-juice”, which by itself did not show catalytic activity on alcoholic fermentation but was able to accelerate this reaction (Harden & Young 1906). Later, Warburg was able to unravel its function in hydride transfer. He also discovered NADPH, which in old literature is referred to as triphosphopyridinucleotide (TPN) or coenzyme II (Warburg & Christian 1936) while NADH was known as coenzyme I or DPN.

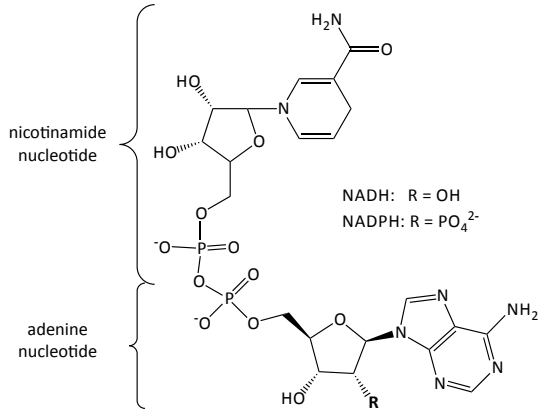


Figure 1-4: Structural formulas of NADH (R = OH) and NADPH (R = PO_4^{2-}).

Both cofactors transport hydride (H^-) ions in their reduced forms NADH and NADPH. This hydride ion is conserved in an energetically rich bond on the nicotinamide ring and can therefore easily be transferred to acceptor molecules. The reduced nicotinamide moiety represents a quinoid system that absorbs light with a wavelength of 340 nm, a characteristic that the oxidized form does not show. The disappearance of the absorption peak at 340 nm due to the consumption of NAD(P)H by an enzyme is an easy way to spectrophotometrically follow the enzyme activity as the decline of NAD(P)H concentration is proportional to the concentration of converted substrate – under the assumption that no uncoupling occurs.

In contrast to the flavin cofactor discussed earlier, which is often tightly (sometimes even covalently) bound by a protein (Valton *et al.* 2008), the nicotinamide coenzymes typically dissociate after each catalytic cycle (Walsh 1979). They are found unbound and in soluble form within the cells acting as electron transport molecules, as no harmful radical is formed during reduction/oxidation of these compounds.

The extra phosphate group of NADPH has no influence on the electron transfer capability of this molecule. Yet, it has an effect on the affinity towards different enzymes. NADPH is typically used as reducing agent in biosynthetic processes, while NAD^+ is responsible for oxidations in the cellular metabolism. Even though the function of the two cofactors is similar, their production and regulation is arranged via different pathways.

This system allows the cell to carefully regulate its energy- and synthesis flow and also spatial separation thereof (Alberts 2008). The redox state of a cell is a clear indicator not only for its metabolic situation but also of the health of the cell (Schafer & Buettner 2001).

While the extra phosphate group of NADPH does not affect the redox properties, it clearly has an effect on the price of the two different nicotinamide coenzymes. The reason for this is the fact that these cofactors have to be isolated from biomaterial. In case of eukaryotic tissue, *e.g.* rat liver, approximately 1 µmol per gram of wet weight NAD^+/NADH can be isolated, outnumbering the $NADP^+$/NADPH concentration by a factor of ten (Reiss *et al.* 1984). This correlates well with the commercial price of the two coenzymes: while NADH is already expensive, also here NADPH surpasses it by one order of magnitude. This makes industrial processes with enzymes that consume NADH and especially NADPH in stoichiometric amounts extremely expensive and therefore often not competitive with chemical methods.

Different solutions have been discussed in the past decades addressing the question of regeneration of these valuable coenzymes in biotechnological processes. The most comfortable solution clearly is the use of whole cells, using the cellular machinery together with the recombinantly expressed enzyme of interest (Stewart 1997). But cells and cell components are not always compatible with the target reaction or process conditions. Then the use of an isolated enzyme is the method of choice. However, an efficient external regeneration system is needed to provide the enzyme with its regenerated coenzyme. Several photo- and electrochemical approaches have been tested so far to replace a second coenzyme regenerating enzyme system, but often with only poor efficiency (van der Donk & Zhao 2003, Hollmann *et al.* 2006, Hollmann *et al.* 2007).

Using a second isolated enzyme for cofactor regeneration seems to be the most suitable method so far – as long as the sacrificial substrate is cheap and no interference with the desired enzymatic reaction (activity and selectivity) takes place. Especially dehydrogenases are widely used for this. A relatively recent promising example is phosphite dehydrogenase (PTDH) from *Pseudomonas stutzeri* WM88 (Costas *et al.* 2001) which merely uses phosphite for regeneration of NADPH or NADH.

In our group we experimented with fusing PTDH with a number of Baeyer-Villiger monooxygenases using an in-house developed expression vector (Torres Pazmiño et al. 2008). In an attempt to optimize this expression system, we have now used a thermostable variant of the PTDH (Johannes et al. 2005, Johannes et al. 2006, Woodyer et al. 2006) and tested the influence of the linker length and composition on the properties of the fusion enzymes. The results of this work can be found in ***chapter 2****.*

In ***chapter 4 & 5*** *we describe a new subclass of class B flavoprotein monooxygenases (Type II FMOs), that – in contrast to other class B flavoproteins - shows no preference for either NADH or NADPH. This is especially interesting when considering the applicability of these newly discovered enzymes for Baeyer-Villiger oxidations.*

1.4. Flavoprotein monooxygenases

Biochemical studies have not only shown that flavoprotein monooxygenases are able to perform a wide range of different oxygenation reactions, they are also able to perform such reactions with highly chemo-, regio- and/or enantioselectivity. A classification for external flavin-dependent monooxygenases was introduced by van Berkel *et al*. in 2006 as they form a major and ubiquitous group of monooxygenases, performing oxygenation reactions that include hydroxylations, epoxidations, oxidations of heteroatoms, halogenations and Baeyer-Villiger oxidations. However, the protein environment and shape of the active site of each monooxygenase is responsible for its specific selectivity and reaction type. This microenvironment on one hand selects the potential substrates and their positioning within the enzyme, on the other hand it modulates the electro- or nucleophilic character of the peroxyflavin (van Berkel *et al.* 2006). Flavoprotein monooxygenases can be divided in six subclasses based on similarity in amino acid sequence and structural homology (Figure 1-5). Of course, all members of the flavoprotein monooxygenases employ a flavin, usually FAD. Only class C monooxygenases utilizes FMN.

Also all of them employ the nicotinamide coenzymes NADH and/or NADPH as electron donor. So far, class B monooxygenases were described using exclusively NADPH. However, by our recent discovery of NADH-accepting class B monooxygenases (**Chapter 4 & 5**), this has to be reconsidered.

All flavoprotein monooxygenases require at least two functionalities: a mono-oxygenase moiety for the oxygenation reaction and a reductase moiety for the supply of electrons for reduction of the flavin. Here two major differences can be found within the 6 classes (Figure 1-5). In the classes C-F, both components are encoded in at least two different genes and expressed as separate enzymes, whereas class A and B monooxygenases have fused both functionalities in one gene/protomer. The class C two-component monooxygenases are represented by Type II Baeyer-Villiger monooxygenases (Taylor & Trudgill 1986, van der Werf 2000), hydroxylases, and luciferases (light emission by oxidation of long aliphatic aldehydes) (Baldwin & Ziegler 1992, Viviani 2002). Class D monooxygenases typically perform oxygenations of aromatic rings (Prieto & Garcia 1994) or non-activated long chain alkanes (Li *et al.* 2008), while class E monooxygenases perform epoxidations (Panke *et al.* 1998), and class F monooxygenases are able to catalyze halogenations (Keller *et al.* 2000). In these two-component systems the flavin acts as a coenzyme, which is reduced by an auxiliary reductase component. The reduced flavin coenzyme displays a high affinity for the oxygenase component, forming an oxygenation catalyst.

In contrast to the above-mentioned flavoprotein monooxygenases, the single component flavoprotein monooxygenases (class A and B) normally have a tightly bound flavin cofactor, acting as prosthetic group. Both classes have at least one dinucleotide binding domain which can easily be identified by the Rossman fold sequence motif GxGxxG. Such a domain is able to bind the ADP moiety of FAD or NAD(P)H (Wierenga *et al*. 1986). Regardless how important tight binding of the flavin cofactor for these enzymes might be, no flavoprotein monooxygenase has yet been found having the flavin covalently tethered to the protein which is in contrast to flavoprotein oxidases (de Jong *et al*. 1992, Mewies *et al*. 1998, Heuts *et al*. 2009).

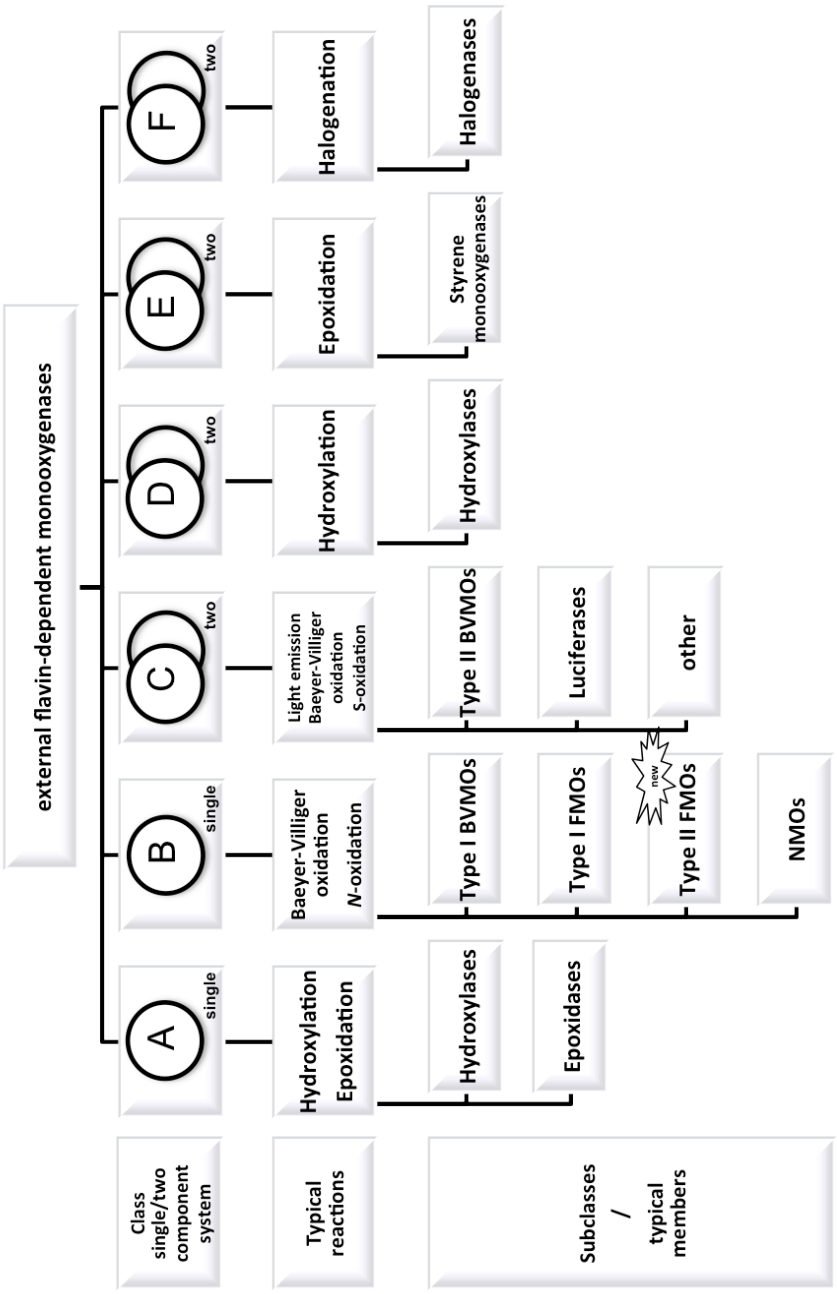


Figure 1-5: Updated classification of external flavoprotein monooxygenases (van Berkel 2006).

Class A monooxygenases have been found to catalyze epoxidation and hydroxylation reactions on aromatic rings. They usually show a narrow substrate specificity, acting on a selected number of aromatic compounds with an activating hydroxyl- or amino group (van Berkel *et al*. 2006). Even though class A enzymes are using both FAD and NAD(P)H, they only have one dinucleotide binding domain and therefore cannot bind both cofactors at the same time. This is in line with the reaction mechanism, as the FAD is bound tightly and the $NAD(P)^+$ is released directly after reduction of the FAD.

In contrast to class A enzymes, class B monooxygenases have two dinucleotide binding domains, each involved in binding one of the cofactors FAD and NAD(P)H. Different to class A monooxygenases, they keep the $NADP^+$ bound during the whole catalytic cycle as it stabilizes the peroxyflavin enzyme intermediate. The coenzyme is released, together with the elimination of water from hydroxyflavin, at the end of the cycle (Figure 1-3, 4 → 5) (van den Heuvel *et al.* 2005, Torres Pazmiño *et al.* 2008).

*The research described in this thesis focused on the discovery and biocatalytic exploration of novel class B flavoprotein monooxygenases (**chapter 3**) with special emphasis on Type I BVMOs. During the project, a new subgroup of class B flavoprotein monooxygenases has been discovered: Type II FMOs (**chapter 4 & 5**).*

1.4.1. Class B flavoprotein monooxygenases

As the members of class B flavoprotein monooxygenases are able to oxidize not only carbon atoms (Baeyer-Villiger oxidations and epoxidations) but also heteroatoms, they are also called the class of "multifunctional flavin-containing monooxygenases". Class B flavoprotein monooxygenases can be divided into different subfamilies based on specific sequence motifs. So far three subfamilies were known and it was found that all members are strictly NADPH dependent: Type I Baeyer-Villiger monooxygenases (BVMOs), flavin-containing monooxygenases (FMOs) and *N*-hydroxylating monooxygenases (NMOs) (Fraaije *et al*. 2002).

*We could proof the existence of at least one additional subclass. Members of this subclass share high sequence homology with members of the above-mentioned subclasses, mainly the FMOs. Yet, their coenzyme dependence sets them apart as they accept both NADH and NADPH, with similar affinity. While referring to the "old" subclass of FMOs as Type I FMOs we named the newly discovered subclass Type II FMOs (**chapter 4**).*

Additionally to the two Rossman fold motifs, BVMOs and Type I FMOs display a typical fingerprint motif that makes them easy to be recognized in databases. In contrast, at the same position the *N*-hydroxylating monooxygenases (NMOs) only share a conserved histidine. So far only a few NMOs from bacteria and fungi have been reported. They are often involved in the biosynthesis of siderophores and they typically convert long-chain primary amines by *N*-hydroxylation (Stehr *et al.* 1998).

The fingerprint motifs of BVMOs and FMOs are quite similar to each other: BVMOs are characterized by the (FxGxxxHxxxWP/D) and FMOs by the (FxGxxxHxxxYK/R) sequence motif (Fraaije *et al.* 2002) and both are found between the two Rossman folds. We could show that the motif in the case of BVMOs allows more variation than thought before. Furthermore, the motif is located in a loop region of BVMO/FMO structures that is not directly involved in catalysis. This hints to a rather poor sequence motif conservation in distant homologs.

*By using the newly identified BVMOs, we determined an additional BVMO-typifying sequence motif involving also active side residues. This motif is only present in BVMOs and clearly distinguishes them from FMOs: [A/G]GxWxxxx[F/Y]P[G/M]xxxD (**chapter 3**) (Riebel et al. 2012).*

BVMOs primarily catalyse the Baeyer-Villiger oxidation of ketones or aldehydes to esters or lactones respectively. They are also able to oxygenate heteroatom-containing

compounds (N, S, B or Se containing compounds). Usually FMOs only catalyse the latter activity.

*With the discovery of Type II FMOs we added a new subclass to the class B flavoprotein monooxygenases (**chapter 4**). Type II FMOs show a slightly modified FMO sequence motif while exhibiting different catalytic properties: they accept both NADH and NADPH with similar affinities and they are able to perform Baeyer-Villiger oxidations.*

1.4.2. Type I Baeyer-Villiger monooxygenases (BVMOs)

Since its discovery more than 100 years ago the Baeyer-Villiger reaction is one of the key reactions in organic chemistry (Baeyer & Villiger 1899). The reaction entails the conversion of carbonylic compounds into the corresponding esters or lactones by the use of peracids. The reaction proceeds via formation of a tetrahedral intermediate, the so-called Criegee intermediate, followed by a concerted migration of one adjacent carbon to the oxygen under the elimination of carboxylic acid (Figure 1-6).

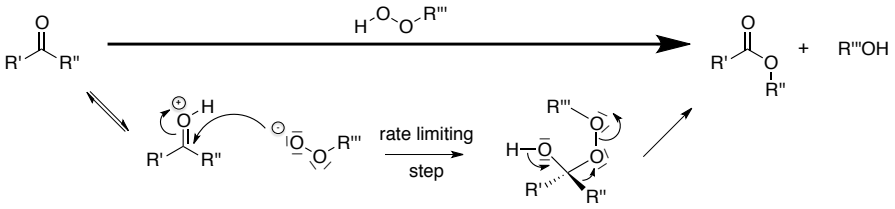


Figure 1-6: Baeyer-Villiger oxidation.

Its key features are the predictable regiochemistry as well as its stereospecificity, as the migrating group retains its configuration. This makes the Baeyer-Villiger reaction a useful tool for asymmetric synthesis with a relative migration aptitude: tertiary alkyl > cyclohexyl > secondary alkyl > benzyl > phenyl > primary alkyl > cyclopentyl, cyclopropyl > methyl (Renz & Meunier 1999).

Already in 1948 Turfitt discovered bioactivity performing this reaction (Turfitt 1948). Yet, 20 years passed before the first BVMO was isolated and characterized (CHMO) (Forney & Markovetz 1969). Malito *et al.* solved the first crystal structure for a member of this class just recently (PAMO), as shown in Figure 1-7 (Malito *et al*. 2004).

Even though a chiral variant of the Baeyer-Villiger reaction is the most prominent reaction catalyzed by BVMOs (for general reviews on Baeyer–Villiger reactions, see: Renz & Meunier 1999, Brink *et al*. 2004, Torres Pazmiño *et al*. 2010; for reviews on stereoselective Baeyer–Villiger oxidations see: Strukul 1998, Blom 2000, Mihovilovic *et al*. 2004, Balke *et al.* 2012), oxygenations of heteroatoms like sulfur, nitrogen, boron, phosphate and others, as well as epoxidation reactions have also been reported (Latham & Walsh 1986, Latham *et al.* 1986, Ottolina *et al.* 1999, Colonna *et al.* 2002, Gonzalo *et al*. 2005). As these reactions are difficult to perform selectively with chemical catalysts, BVMOs are clearly powerful tools for synthetic chemistry.

Various studies have been performed with the aim to identify the responsible residues for the different substrate specificities of BVMOs. As most of the known BVMOs are rather unstable, the best model enzyme so far turned out to be the thermostable phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* (Fraaije *et al.* 2005). It shows a remarkable stability ($t_{1/2}$ = 24 h at 52 °C) and for a long time it was the only BVMO with a known crystal structure, making valuable structural data available (Figure 1-7) . Several engineered PAMO mutants with altered substrate specificities have been created and allow a deeper insight into important residues for selectivity and substrate recognition (Dudek *et al.* 2011). But even by using this knowledge it is still impossible to rationally redesign a biocatalyst with predefined features. Nevertheless, the interest in effective and selective oxidative biocatalysts is high in academia and industry. Therefore, it is tempting to exploit the available typical sequence motifs for genome-mining approaches aimed at the discovery of new BVMOs displaying novel or improved selectivities.

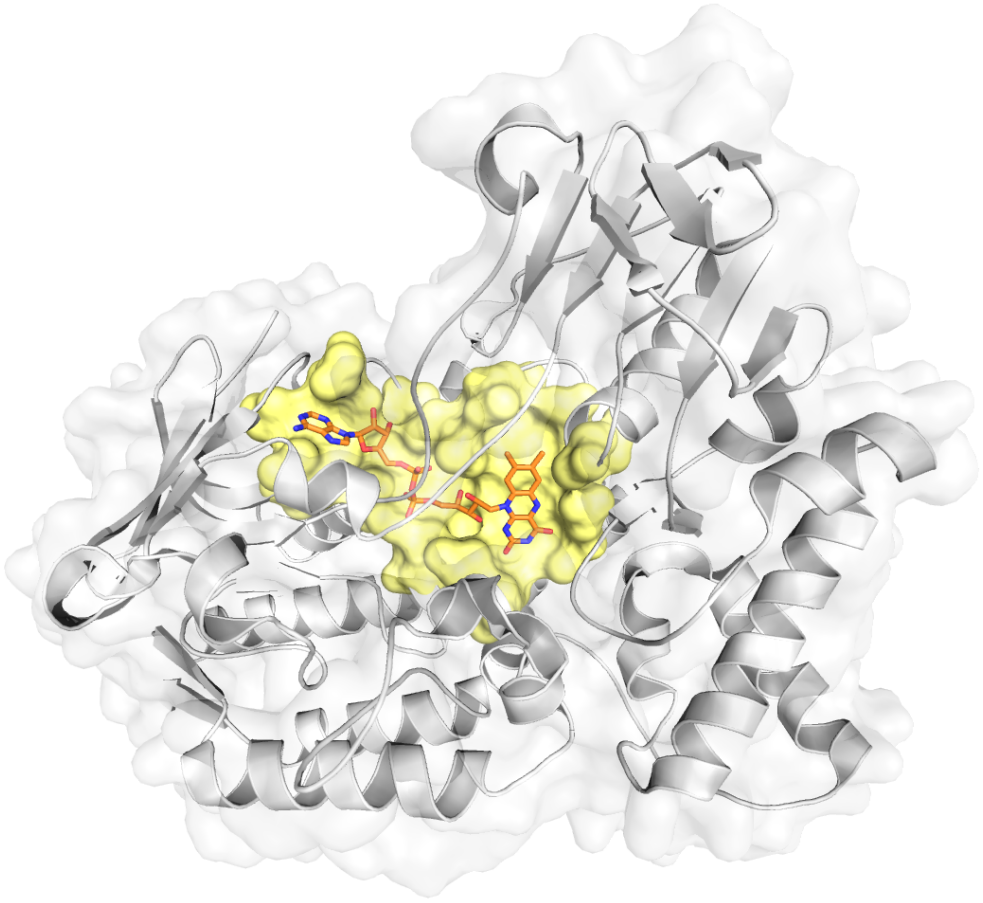


Figure 1-7: Crystal structure of phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* (pdb: 1W4X). The FAD cofactor is shown as orange sticks with its binding site in yellow.

*We were able to identify, clone, recombinantly express and characterize 22 new Type I BVMOs using the genome sequence of Rhodococcus jostii RHA1 (**chapter 3**). We were also able to identify a new class B flavoprotein monooxygenase subclass, Type II FMOs, that can also act as BVMOs (**chapter 4 & 5**).*

1.5. *Rhodococcus jostii* RHA1

Rhodococcus is a widely occuring genus of the *Actinomycetales*, an order of non-motile Gram-positive bacteria. With only a few exceptions, *Rhodococcus* species are nonpathogenic aerobic soil bacteria. *Rhodococcus jostii* RHA1 was first isolated from contaminated soil in 1995 by virtue of its ability to degrade polychlorinated biphenyls (Seto *et al.* 1995). Its genome was sequenced in 2006 and with 9.7 Mb this organism has one of the largest bacterial genomes sequenced to date (McLeod *et al*. 2006). The genomic arrangement is divided into one linear chromosome and three linear plasmids. The organism itself displays a high catabolic diversity that makes it interesting for environmental and biotechnological purposes. Other interests lay in engineering this organism towards an alternative expression and production host (van der Geize & Dijkhuizen 2004).

1.6. Aim of this thesis

This thesis was performed within the BIOMOX cluster of the "Integrated Biosynthesis and Organic Synthesis (IBOS)" programme of the Netherlands Organization for Scientific Research (NWO). The aim of this project was the development of new biocatalytic platforms performing enantio- and/or regioselective oxidation reactions. The targeted reactions of the project were hydroxylations, Baeyer-Villiger oxidations, epoxidations and sulfoxidations. *Rhodococcus jostii* RHA1 was chosen as source for new biocatalysts. In total four academic groups were actively involved in this research project. The focus in the Fraaije research group was on the discovery of novel flavoprotein monooxygenases catalyzing Baeyer-Villiger oxidations and sulfoxidations.

As interesting as flavoprotein monooxygenases may be, there are still some drawbacks in their applicability. One is the fact that they are cofactor dependent and for the use in an industrial process an NAD(P)H regeneration system is required. In **chapter 2** we have addressed this problem by creating and optimizing a fused coenzyme regeneration system (pCRE2), resulting in the effective production of self-sufficient monooxygenases.

For the work described in this thesis we decided to explore the "oxidative treasure" in the genome of *Rhodococcus jostii* RHA1 (McLeod *et al.* 2006). Even though BVMOs are widespread in bacteria and fungi, the average distribution is only around one or two BVMO-encoding genes per microbial genome (de Gonzalo *et al.* 2010). FMOs and NMOs are even more rare in microbes. An unusually high number and variety of oxidative enzymes could be identified in the predicted proteome of *R. jostii*. We were able to identify, express and characterize not only 22 new BVMOs (**chapter 3**), but also seven new FMOs and one putative NMO (**chapter 4**) from this bacterium.

The newly discovered BVMOs (**chapter 3**) were characterized using 39 different potential substrates in a newly developed activity-screening assay. For this approach, an advanced method to determine the concentration of NADPH-dependent enzymes in cell extracts was developed. Additionally, with this large set of enzymes we postulated a new BVMO-identifying fingerprint motif, which can facilitate the future identification of such enzymes in genome databases.

The flavoprotein monooxygenases described in **chapter 4** formed a separate clade of sequence-related proteins being distinct from the known FMOs, NMOs, and BVMOs. They also showed distinct enzyme characteristics, as they accept both NADH and NADPH with similar affinities. This clearly separates them from the other class B monooxygenase subclasses and we therefore proposed a new subclass of class B flavoproteins: Type II FMOs.

Surprisingly, three of the newly identified Type II FMOs were also able to catalyze Baeyer-Villiger oxidation reactions. These three enzymes also share a high sequence identity with each other, including an N-terminal extension of about 160 residues. These three enzymes have been intensely characterized in **chapter 5**. Stopped-flow kinetic analysis proofed that the reaction mechanism indeed follows real BVMO characteristics. A careful inspection of the active site residues between Type I BVMOs, Type I FMOs and the three special members of the Type II FMOs revealed that two conserved residues in the newly identified class II FMOs represent a completely new active site architecture around the flavin. With mutagenesis studies we addressed potentially important active site

residues. We could show that replacing the targeted residues has a huge effect on the ability to efficiently bind the flavin cofactor.

The **chapters 2 – 5** describe the results of a project aimed at the discovery of new oxidative biocatalysts. The work has resulted in the recombinant production and biocatalytic characterization of >30 new flavoprotein monooxygenases. With this set, a wide range of reactions can be performed. **Chapter 6** provides an overview of the obtained results.

Any sufficiently advanced technology is indistinguishable from magic.

Arthur C. Clarke

CHAPTER 2

EFFICIENT BIOOXIDATIONS CATALYZED BY A NEW GENERATION OF SELF-SUFFICIENT BAEYER-VILLIGER MONOOXYGENASES

This chapter is based on: ***ChemBioChem (2009), 10, 2595-2598*** *by Daniel E. Torres Pazmiño, Anette Riebel, Jon de Lange, Florian Rudroff, Marko D. Mihovilovic and Marco W. Fraaije.*

Abstract: To facilitate the purification and application of NADPH-dependent BVMOs we have developed an expression system for the production of bifunctional biocatalysts: BVMOs fused to phosphite dehydrogenase (PTDH). By fusing BVMOs to PTDH we have created self-sufficient enzymes that only require phosphite for regenerating a catalytic amount of NADPH. In this chapter efforts to optimize the expression system are discussed.

2.1. Introduction

Over the last decades industrial interest in oxidative biocatalysis has increased significantly. One of the most prominent enzyme families that catalyze a variety of different oxidations are Baeyer-Villiger monooxygenases (BVMOs). These biocatalysts are part of an exclusive family of flavin-containing enzymes that catalyze the biotransformation of aldehydes and (a)cyclic ketones to their corresponding esters and lactones (Kamerbeek *et al.* 2003, van Berkel *et al*. 2006, Torres Pazmiño & Fraaije 2007, Alphand *et al.* 2009, Kayser 2009). Additionally, these enzymes are also known for their capability to oxidize heteroatoms (sulphur, nitrogen, boron) and perform epoxidation reactions (Latham & Walsh 1986, Latham *et al.* 1986, Ottolina *et al.* 1999, Colonna *et al.* 2002, Gonzalo *et al.* 2005, Rial *et al.* 2008). Type I BVMOs utilize flavin adenine dinucleotide (FAD) as cofactor and NADPH as electron donor in order to activate molecular oxygen (O_2) and generate a reactive C4a-peroxyflavin intermediate. This enzyme intermediate acts similar to an organic peracid and reacts with the organic substrate resulting in formation of the oxygenated product (Ryerson *et al.* 1982, Sheng *et al.* 2001, Torres Pazmiño *et al*. 2008). BVMOs have been shown to carry out these oxidative reactions in a highly regio-, stereo- and enantioselective manner, indicating that these enzymes are interesting candidates for various biocatalytic applications (Walsh & Chen 1988, Roberts & Wan 1998, Mihovilovic *et al.* 2002, Mihovilovic 2006). An obstacle of using BVMOs in a cost efficient way is the requirement of stoichiometric amounts of expensive NADPH coenzyme. Several coenzyme regeneration methods have been explored in the recent years (van der Donk & Zhao 2003, Wichmann & Vasic-Racki 2005, Hollmann *et al*. 2006). The most efficient approach is based on the regeneration of NADPH by using a two-enzyme system (either as isolated enzyme or in whole cells) (Stewart 1997, Zambianchi *et al.* 2002, Schulz *et al.* 2005, Lee *et al*. 2007).

Typically, $NAD(P)^+$-dependent dehydrogenases are applied to regenerate the reduced coenzyme at the expense of a cheaper sacrificial substrate *(*Peretz *et al.* 1997, Tishkov *et al.* 1999, Tishkov & Popov 2006). Phosphite dehydrogenase (PTDH, EC 1.20.1.1) is an excellent example of such NAD(P)H regenerating enzymes. This enzyme catalyzes the

thermodynamically driven and therefore almost irreversible oxidation of (cheap) phosphite to phosphate by simultaneous reduction of (expensive) $NAD(P)^+$ (Vrtis *et al.* 2002, van der Donk & Zhao 2003). In our research group we have recently investigated the utility of PTDH in combination with BVMOs and could show, that in this combination none of the two individual reactions is disturbed by any side-reactions or interferences with either substrates or products. By covalently fusing the PTDH to the N-terminus of various BVMOs, so-called CRE/BVMOs were created (CRE; coenzyme regenerating enzyme) (Torres Pazmiño *et al*. 2008). By fusing the BVMOs to PTDH the expression levels of the targeted BVMO increased and the overall solubility was improved when compared with expression of the non-fused BVMOs. This indicates that the PTDH-gene is also functioning as a very effective expression tag.

To summarize, the advantages of these bi-functional biocatalysts are that with (*i*) minimum effort in terms of enzyme production and isolation and (*ii*) maximum simplicity with respect to coenzyme regeneration, challenging enzymatic oxidations can be performed.

The first generation of these self-sufficient CRE/BVMOs were successfully applied for the bioconversions of a wide variety of substrates by using whole cells, cleared cell extracts (CCEs) and purified enzyme. Interestingly, biooxidations using cleared cellextracts could be performed without addition of $NADPH/NADP^+$. This can be explained by the fact that sufficient amounts of coenzyme are released for catalysis when the *Escherichia coli* cells are ruptured.

However, the activity of the first reported self-sufficient BVMOs decreased significantly over time due to instability of the PTDH fusion partner. By a successful combination of site-directed and random mutagenesis approaches, Zhao and coworkers have been able to improve the thermostability of PTDH, as well as its resistance towards organic solvents. In comparison to the wild-type, this PTDH mutant (18x PTDH) contained 18 mutations and showed a 7000-fold longer life time at 45 °C (Johannes *et al.* 2005, Johannes *et al.* 2006, Woodyer *et al.* 2006).

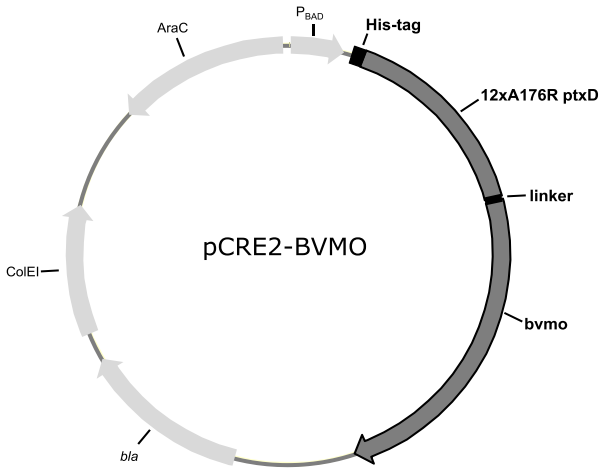


Figure 2-1: plasmid map of the expression vector pCRE2/BVMO.

In this chapter we present fusion engineering and biocatalytic studies of a new generation of self-sufficient BVMOs (Figure 2-1). Except for replacing the original PTDH by a more robust variant, we also investigated the influence of the linker length and composition between both fusion partners. In comparison to the previous generation, these second generation self-sufficient CRE2/BVMOs posses (i) a thermostable variant of PTDH as fusion partner for effective coenzyme regeneration, (ii) a codon-optimized PTDH mutant gene for excellent expression in E. coli (Villalobos et al. 2006), and (iii) a histidine purification tag at the free N-terminus of the PTDH partner. In this way, the dehydrogenase fusion partner is not only exploited for its ability to efficiently regenerate nicotinamide coenzymes, but also as expression and purification tag (Thapa et al. 2008).

2.3. Results and Discussion

2.3.1. Expression of CRE2/BVMOs and characterization of kinetic parameters and thermostability

For this study, six representative BVMOs were selected and fused to the C-terminus of the thermostable PTDH variant; cyclopentanone monooxygenase from *Comamonas* sp. strain NCIMB 9872 (CPMO, EC 1.14.13.16)(Griffin & Trudgill 1976, Iwaki *et al.* 2002, van Beilen *et al.* 2003), cyclohexanone monooxygenase from *Acinetobacter* sp. NCIMB 9871 (CHMO, EC 1.14.13.22) (Donoghue *et al.* 1976), 4-hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB (HAPMO, EC 1.14.13.81) (Kamerbeek *et al*. 2001), ethionamide monooxygenase from *Mycobacterium tuberculosis* H37Rv (EtaA, EC 1.14.13.92) (Fraaije 2003), phenylacetone monooxygenase from *Thermobifida fusca* (PAMO; EC 1.14.13.92) and the M446G PAMO mutant (Fraaije *et al.* 2005, Torres Pazmiño *et al.* 2007). Expression levels of all six CRE2/BVMOs in *E. coli* TOP10 were excellent and subsequent purification was performed in a one step process by using a Ni-NTA column. Even CRE2−CPMO was obtained in good purity while previous attempts to purify CPMO from a recombinant host failed, tentatively due to stability reasons (Iwaki *et al.* 2002, van Beilen *et al.* 2003).

The activities and kinetic parameters of the isolated CRE2/BVMOs were determined using the primary substrates of each enzyme. In general, the catalytic activities (k_{cat}) and Michaelis-Menten constants (K_M) of the BVMOs fused to the thermostable PTDH were found to be very similar to those of the non-fused enzymes (Table 2-1). The catalytic efficiency (k_{cat}/K_M) of the PTDH subunit for $NADP^+$ was found to vary between 93 and 1300 $mM^{-1}\ s^{-1}$, depending on its fusion partner. The variance in these values may reflect the level of flavin cofactor incorporation in each fused enzyme and are at the same order of magnitude compared to the non-fused PTDH variant (k_{cat}/K_M = 2.5×10^2 $mM^{-1}\ s^{-1}$) (Johannes *et al.* 2006)

Table 2-1: Kinetic parameters of 2nd generation self-sufficient BVMOs (CRE2/BVMOs).

	BVMO subunit			PTDH subunit
CRE2-BVMO	k_{cat} (s^{-1})	$K_{M,\ ketone}$ (μM)	$K_{M,\ NADPH}$ (μM)	$k_{cat}/K_{M,NADP+}$ ($mM^{-1}\ s^{-1}$)
CRE2–PAMO	**2,6** (phenylacetone)	**60**	**3,5**	**$3,6 \times 10^2$**
CRE2–CHMO	**13** (cyclohexanone)	**5,5**	**4**	**$9,3 \times 10^1$**
CRE2–CPMO	**1,8** (cyclopentanone)	**< 1,0**	**< 1,0**	**$1,8 \times 10^2$**
CRE2–HAPMO	**12** (4-hydroxyacetophenone)	**7,8**	**8,3**	**$6,1 \times 10^2$**
CRE2–EtaA	**$1,0 \times 10^{-1}$** (phenylacetone)	**$1,3 \times 10^3$**	**> $1,2 \times 10^2$**	**$6,9 \times 10^2$**
CRE2–$PAMO_{M446G}$	**1,8** (benzaldehyde)	**$2,5 \times 10^3$**	**< 2,0**	**$1,3 \times 10^3$**

In order to determine the stability of these new CRE2/BVMOs, activity measurements were carried out at elevated temperatures. Incubation of CRE2–CHMO at 37°C resulted in a decrease in BVMO activity over time ($t_{½,}$ 37 °C = 95 minutes), whereas no loss in activity was observed for the PTDH subunit during 24 hours. This clearly indicates that the activity and stability of the dehydrogenase is not affected by the thermal inactivation of its fusion partner. Incubation of CRE2–PAMO at 50 °C showed that the PTDH subunit is also suitable for coenzyme regeneration at elevated temperatures, showing a half-life time ($t_{½,}$ 50°C) of 39 hours. Additionally, we determined that the PTDH subunit tolerates methanol as organic cosolvent. The initial specific activity of the dehydrogenase subunit of CRE2–PAMO was found to be hardly affected when using up to 40 % methanol. In the presence of 30 % methanol at 30 °C, a slow decrease in dehydrogenase activity was observed ($t_{½}$ = 2 hours).

2.3.2. Linker studies

Even though it has been shown that fusing a BVMO to PTDH typically has almost no large influence on the performance of the BVMO activity, still some small effects on catalytic performance have been observed (Torres Pazmiño *et al.* 2008). In the case of CRE2-PAMO and CRE2-$PAMO_{M446G}$ the kinetic parameters for the BVMO are rather unaffected, while

the PTDH appears to suffer from the fusion (Table 2-2). While the k_{cat} values increase for the fused PTDH, the K_M values increase significantly, resulting in similar catalytic efficiencies (k_{cat}/K_M).

Table 2-2: Kinetic parameters of 2nd generation self-sufficient CRE2/PAMO in comparison to non-fused enzymes.

	$k_{cat, BVMO}$ (s^{-1})	$K_{M, ketone}$ (μM)	$K_{M, NADPH}$ (μM)	$k_{cat, PTDH}$ (s^{-1})	$K_{M, phosphite}$ (μM)	$K_{M, NADP+}$ (μM)
CRE2– PAMO / PAMO	**2,6** / 3,1	**60** / 80	**3,5** / 0,7	**6,2** /1,4	**255** / 36	**17 /** 5,5
CRE2–PAMO$_{M446G}$ / PAMO$_{M446G}$	**1,8** / 2,9	**2500** / 1600	**< 2**	**10,5** /1,4	**340** / 36	**8,6** / 5,5

Kinetic parameters of purified CRE2-BVMOs and non-fused BVMOs at 25°C. BVMO activity for PAMO were measured using phenylacetone as substrate while PAMO$_{M446G}$ activity was measured using benzaldehyde. The parameters of the non-fused BVMOs were taken from literature (Torres Pazmiño *et al*. 2007 & 2008).

A possible explanation for these observations might be incomplete FAD incorporation in the fusion-versions, due to the overall protein folding. As the protein concentration determination is done via analysis of the incorporated FAD in the BVMO partner, the concentration of active PTDH might be underestimated. Additionally, wild type PTDH naturally forms dimers and experiments with the previous fusion system (not thermostable PTDH) gave indications that fusion did not prevent dimerization. As an attempt to improve the FAD binding, we aimed to optimize the linker length and constellation to allow a better folding and cofactor binding.

Length and composition of peptide linkers are known to affect protein stability and domain-domain interactions of fusion enzymes (Argos 1990, Robinson & Sauer 1998, Evers *et al.* 2006). All CRE2/BVMOs described above contain a peptide linker consisting of six amino acids, *i.e.* Ser-Arg-Ser-Ala-Ala-Gly. In order to determine whether the linker has an effect on the stability or performance of the individual enzymes in the fused bifunctional enzyme, we constructed three additional variants of CRE2−PAMO with different linkers, based on linker properties of natural fusion proteins and biochemical studies on fusion proteins.

In nature fusion enzymes are common. Based on the study of Argos in 1990 a natural linker normally fulfills several criteria: (*i*) it is composed of small, non-bulky and polar

amino acids (aa) with a maximum of two charges, (*ii*) it spans a minimum of 5 aa to a maximum of 20 aa and is organized in an extended coil structure with mainly glycincs, and (*iii*) it may include turns by the insertion of prolines and asparagines. Table 2-3 shows a list of preferred and avoided aa in natural linkers.

Table 2-3: favored and unfavored amino acids in natural linkers (Argos, 1990).

strongly preferred	Gly, Ala, Ser, Thr
preferred	Asp, Gln, Lys, Asn, Pro
disfavored	Met, Cys, Ile, Tyr, Phe, Trp, Leu, Val
strongly disfavored	Arg, Glu

We set out to test several linker variants (Table 2-4). Variant 1 represents a linker-sequence that is often used in artificial fusion enzymes (Robinson & Sauer 1998, Evers *et al.* 2006), consisting of glycines and serines. This linker results in a random coil structure ("glycine-rich"). Variant 2 consists of all preferred amino acids that can be found in natural fusion enzymes ("random"). As a last variant (3) we reduced the linker length to a minimum of 1 amino acid, tryptophane, representing an amino acid that nature tends to avoid ("short").

Table 2-4: length and peptide sequences of the different linker versions.

linker variant	peptide sequence	length (aa)
original fusion	SRSAAG	6
variant 1 (glycine-rich)	SSGGSGGSGGSAG	13
variant 2 (random)	SSATGSATGSAG	11
variant 3 (short)	W	1

When expressing the different linker variants of CRE2−PAMO, no difference in expression levels could be detected. All fusion proteins show the same massive overexpression when compared with the original CRE2-PAMO. The thermostability of the variants was tested in two different ways: using the ThermoFAD method (Forneris *et al.* 2009) and by monitoring the activity after incubation at elevated temperatures. By measuring the melting

temperature (T_M) via the Thermo FAD method no significant difference could be detected (Figure 2-2). All CRE2-PAMOs (variants and original construct) showed a melting temperature of 58.5°C.

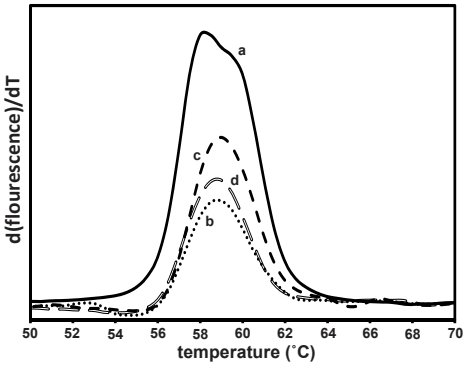


Figure 2-2: Thermo FAD measurements for the determination of the melting temperature with (a) original linker, (b) glycine-rich linker, (c) random linker, (d) short linker.

Also, when monitoring the remaining activity of both subunits of the CRE2−PAMO variants in time, all four variants were found to have similar thermostabilities (Figure 2-3). The glycine-rich linker was found to result in the most robust fusion protein. These results indicate that the termini of both enzymes are sufficient flexible and solvent-exposed to yield an effective and thermostable fusion enzyme, regardless of the nature of the peptide linker between these two subunits.

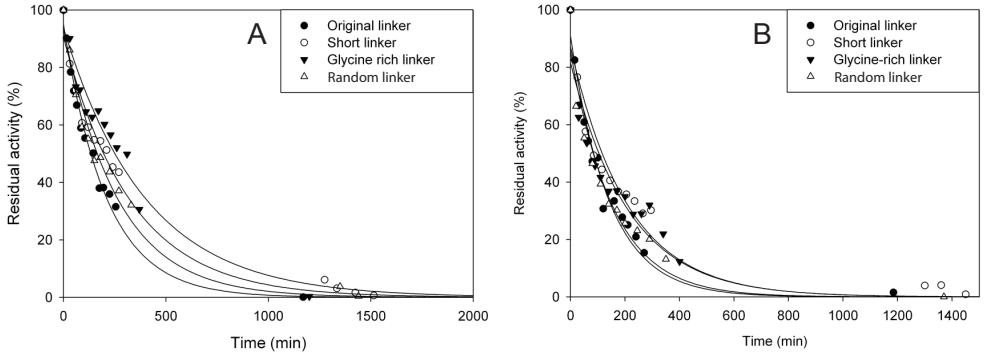


Figure 2-3: Residual activity of the PAMO (A) and PTDH (B) subunit of CRE2−PAMO containing different linker variants at 55 °C, followed in time and fitted with an exponential decay (y = a*e –x/b). The following lifetimes (t½, 55 °C) were determined for these variants: A (PAMO) original linker, 208 min; glycine-rich linker, 400 min; random linker, 264 min; short linker, 333 min. B (PTDH) original linker, 142 min; glycine-rich linker, 212 min; random linker, 156 min; short linker, 208 min.

While no drastic effects on stability were observed, a larger effect on FAD incorporation was found. While the original PTDH-PAMO is occupied for 67% by FAD, in all other linker variants this percentage was < 50% (Table 2-5). This indicates that the original linker is beneficial for FAD incorporation in PAMO.

Table 2-5: FAD incorporation in the CRE2-PAMO linker variants.

linker variant	FAD incorporation: % apo PAMO
original fusion	33
variant 1 (glycine-rich)	54
variant 2 (random)	52
variant 3 (short)	59

Based on the results above we conclude that the type of linker does not have a huge effect on the stability behaviour of the fused enzyme. The original linker outperforms the alternative linkers with respect to cofactor binding, even though it contains one amino acid that would have been avoided in terms of natural selection. Therefore we decided to continue also future research with the original version of the vector pCRE2.

2.3.3. Biocatalysis

Previously, we demonstrated that bifunctional CRE/BVMOs can be also applied for bioconversions using whole cells and cleared cell extracts. Cleared cell extract containing CRE−CHMO was used to catalyze the oxidation of 4-methylcyclohexanone (**1**) to the corresponding lactone **2** in the presence of phosphite. The biotransformation stopped after a reaction time of 6 hours, while less than 40 % conversion was achieved (>10 mM substrate concentration) (Torres Pazmiño *et al.* 2008). Performing the same transformation with the newly created thermostable CRE2−CHMO, resulted in a more effective oxidation of the substrate (Figure 2-4).

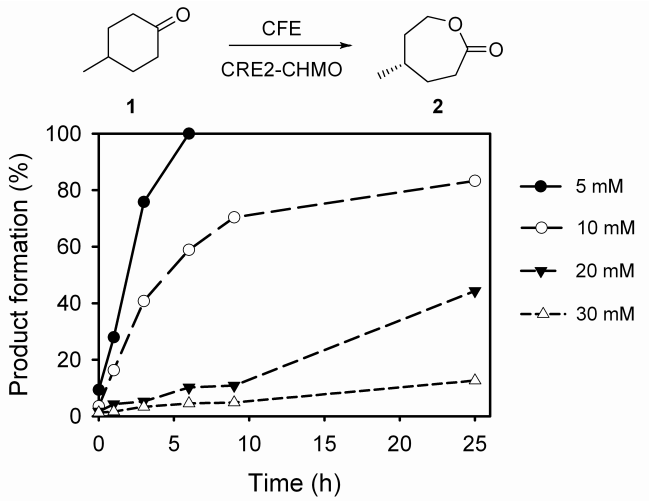


Figure 2-4: Baeyer-Villiger oxidation of 4-methylcyclohexanone 1 by CCE of *E. coli* TOP10 expressing CRE2−CHMO in the presence of 100 mM phosphite at 24 °C, while no coenzyme was added.

In contrast to the first generation self-sufficient BVMOs, the thermostable variant CRE2−CHMO showed no inactivation within the presented time frame. This resulted in the full conversion of 5 and 10 mM of ketone **1** within 24 hours, yielding > 80 % product **2**. Despite possible substrate inhibition, 13 % of conversion was achieved by using 30 mM of ketone **1**, while hardly any conversion was observed when CHMO was fused to the relatively thermolabile PTDH (Torres Pazmiño *et al*. 2008). These results confirm that the instability of PTDH in the first reported self-sufficient CRE/BVMOs was the limiting factor for efficient bioconversions.

Additionally, we investigated the biocatalytic activities of other CRE2/BVMOs in cleared cell extract using bicyclo[3.2.0]hept-2-en-6-one **3** as a model substrate. At 24 °C, 5 mM of bicyclic ketone **3** was efficiently oxidized to the corresponding lactones **4** and **5** (Figure 2-5) by both CRE2−CHMO (100%) and CRE2−CPMO (85%). Only 24 % of product formation was observed with CRE2−PAMO after 24 hours. At higher substrate concentrations (20 mM), CRE2−CHMO fully converted the bicyclic ketone **3** within 24 hours. On the other hand, hardly any conversion was observed with CRE2−CPMO, indicating that the CPMO subunit suffers from substrate inhibition. The oxidation rate of

ketone **3** by CRE2−PAMO was found to increase proportional to the substrate concentration, as similar conversion percentages were obtained using 5 and 20 mM of substrate. This latter observation is in agreement with the low affinity of PAMO towards this bicyclic compound (K_M = 15 mM) (Fraaije *et al.* 2005).

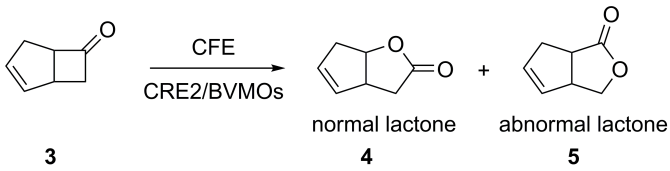


Figure 2-5: Baeyer-Villiger oxidation of bicyclo[3.2.0]hept-2-en-6-one 3.

Moreover, we studied the biooxidation of bicyclo[3.2.0]hept-2-en-6-one at elevated temperatures. In order to stabilize the BVMO subunit to some extent, we supplemented the reaction mixture with 10 µM FAD. Despite the moderate thermostability of the CHMO subunit ($t_{½}$, 37 °C = 95 minutes), CRE2−CHMO was still able to fully convert 5 mM of bicyclic ketone within 3 hours at 37 °C. However, at higher temperature (50 °C) the bifunctional enzyme was inactivated within 20 minutes resulting in a rather poor conversion (21%). CRE2−PAMO showed a higher catalytic activity at elevated temperatures compared to the reaction performed at 24 °C. Even at 50 °C the fusion enzyme was still active, indicating that regeneration of NADPH is possible under these conditions. Unfortunately, hardly any conversion was observed using CRE2−CPMO at 37 °C and 50 °C. This is in agreement with the literature which indicates that CPMO is a labile biocatalyst (Iwaki *et al.* 2002, van Beilen *et al.* 2003).

In this chapter we report on the design, production and evaluation of a new generation of self-sufficient Baeyer-Villiger monooxygenases: CRE2/BVMOs. In comparison to the previous generation, these newly designed bifunctional enzymes contain a multiple mutant of PTDH which has been shown to display a high thermostability (Johannes et al. 2006). In addition, the gene of this PTDH mutant was codon-optimized for expression in E. coli. Furthermore, CRE2/BVMOs contain a

polyhistidine affinity-tag at their N-terminus that facilitates purification of all bifunctional biocatalysts. The purified self-sufficient BVMOs displayed similar kinetic parameters as the non-fused enzymes, suggesting that the properties of these enzymes are hardly affected upon fusion. As expected, their PTDH subunit showed higher (thermo)stability than the dehydrogenase subunit of the first reported self-sufficient BVMOs. Furthermore, these new bifunctional enzymes could be applied as cleared cellextracts for conversion of several ketones without loss of activity. Even at elevated temperatures, sufficient regeneration of NADPH took place for the enzyme-catalyzed Baeyer-Villiger oxidation to occur. Based on these results we can conclude that with the thermostable variant of PTDH as fusion-CRE partner, stable and effective self-sufficient BVMOs can be created that have the potential to be applied in biocatalytic processes.

2.4. Experimental Section

General Methods

Unless otherwise noted, all chemicals and enzymes were obtained from ACROS Organics, Sigma-Aldrich, Jülich Fine Chemicals, Riedel-de Haën Fine Chemicals, Roche Applied Sciences, New England Biolabs and Finnzymes and used without further purification. Oligonucleotide primers were obtained from Sigma Genosys. DNA sequencing was done at GATC (Konstanz). Synthesis of the 18x *ptxD* gene was performed by DNA 2.0. The synthetic gene encoding the 18-fold mutant of phosphite dehydrogenase from *Pseudomonas stutzeri* WM88 was designed using Gene Designer (Villalobos *et al*. 2006). The nucleotide sequence of this mutant was optimized using this software tool for optimal codon usage in *E. coli*. In addition, a 6x histidine tag was added upstream of the synthetic gene to facilitate enzyme purification. Furthermore, recognition sites for the restriction enzymes *Nco*I, *Nde*I and *Xho*I were introduced to facilitate cloning, whereas *Pvu*II and *Hind*III restriction sites were avoided.

Cloning Strategy

For the creation of pCRE2−PAMO, a modified pBAD vector was used in which the *Nde*I site was replaced by the original *Nco*I site. Subsequent cloning of the gene encoding the 18x mutant of phosphite dehydrogenase with N-terminal histidine-tag (from pJ36/18x PTDH) and the *pamO* gene using *Nco*I/*Xho*I, respectively *Pvu*II/*Hind*III, yielded expression vector pCRE2−PAMO. Expression vectors pCRE2−CHMO, pCRE2−CPMO, pCRE2−HAPMO and pCRE2−EtaA were created by subcloning respectively the *chnB*, *cpnB*, *hapE* and *etaA* genes (Donoghue *et al.* 1976, Griffin & Trudgill 1976, Kamerbeek *et al.* 2001, Iwaki *et al*. 2002, Fraaije 2003, van Beilen *et al.* 2003). Expression vector pCRE2−$PAMO_{M446G}$ was created using the QuikChange® Site-Directed Mutagenesis Kit of Stratagene with pCRE2−PAMO as template and following the recommendations of the manufacturer.

Cloning Strategy CRE2−PAMO linker variants

For the creation of pCRE2−PAMO linker variants 1 and 2, the original linker was removed using *Xho*I and *Pvu*II restriction sites. The linearized vector was subsequently treated with Antarctica Phosphatase (New England Biolabs) according to the manufacturer's guidelines. The linker fragments were created using 5'-phosphorylated complementary primers (Table 2-6). These primers were dimerized in a thermocycler by incubating 100 µM of each primer in 10 mM Tris/HCl (pH 8,5) at 72 °C for 2 min, followed by a slow cooling down to 23 °C in 45 minutes. The resulting double-stranded DNA was then ligated to the linearized vector using T4 DNA ligase (New England Biolabs).

Table 2-6: Primer sequences for linker variants 1 & 2.

Linker variant	Primer sequence (5' → 3')
Variant 1 (glycine-rich)	P1-tcgagtggtggctctggtgggagcggtggctcag
	P2-ctgagccaccgctcccaccagagccaccac
Variant 2 (random)	P3-tcgagtgccaccggtagcgcgacgggctcag
	P4-ctgagcccgtcgcgctaccggtgccac

pCRE2−PAMO linker variant 3 was created by removing the PTDH sequence using *NdeI* and *Pvu*II restriction sites. The PTDH gene was then newly amplified with primers including a *NdeI* restriction site in the forward primer. The PCR product was digested with *NdeI* (New England Biolabs) according to the manufacturer's guidelines to create the corresponding overlapping region at the 5'end, while the 3'end stayed blunt. Ligation of the linearized vector and the PCR product was then followed by a QuikChange reaction to remove one nucleotide (Table 2-7).

Table 2-7: Primer sequences for linker variant 3 and quikchange primers sequences.

Variant 3 (short)	Primer sequence
PCR fwd	GGCAGCCATATGCTGCCG
PCR rev	TCAGCCGCAGGGTTGGC
Quik change fwd	GCCAACCCTGCGGCTGACTGGATGGCCGGGCAGACG
Quik change rev	CGTCTGCCCGGCCATCCAGTCAGCCGCAGGGTTGGC

Overexpression and fusion enzyme purification

All self-sufficient BVMOs, as well as the linker variants, were overexpressed in *E. coli* TOP10 using Terrific Broth (TB) medium containing 50 µg mL^{-1} ampicillin and 0,02 % (wv^{-1}) L-arabinose. CRE2−PAMO (and the linker variants), CRE2−$PAMO_{M446G}$ and CRE2−CHMO were overexpressed at 24 °C during 36 hours. CRE2−CPMO, CRE2−HAPMO and CRE2−EtaA were overexpressed at 17 °C during 48 hours. Thereafter, cells were harvested, washed and resuspended in 50 mM Tris/HCl, pH 7,5 containing 10 µM FAD. Cells expressing CRE2−CHMO were resuspended in 50 mM Tris/HCl, pH 7,5 supplemented with 0,5 mM DTT and 10 % (wv^{-1}) glycerol, whereas cells expressing CRE2−EtaA were resuspended in 50 mM Tris/HCl, pH 7,5 supplemented with 1 mM β-mercaptoethanol, 1 mM NaN_3, 100 mM KCl, 10 % (wv^{-1}) glycerol and 1 % (vv^{-1}) Triton X-100. All further steps were carried out at 4 °C. After breaking the cells by ultrasonication and removing cell debris by centrifugation, fusion enzymes were purified using 2 mL Ni-NTA column. After purification, excess of imidazole was removed using an Econo-PAC 10DG desalting column (BioRad). The fusion enzymes were concentrated by ultrafiltration and stored at −80 °C.

Determination of kinetic parameters of isolated fusion enzymes

Activities of purified enzymes were determined spectrophotometrically by monitoring increase or decrease of NADPH in time at 340 nm (ε_{NADPH} = 6,2 mM^{-1} cm^{-1}) or 370 nm (ε_{NADPH} = 2,7 mM^{-1} cm^{-1}). The reaction mixture (1 mL) typically contained 50 mM Tris/HCl, pH 7,5, 100 µM coenzyme, 2 mM ketone or 5 mM Na_2HPO_3 and 0,05 – 1 µM enzyme at 25 °C. The following ketones were used: phenylacetone for CRE2−PAMO and CRE2-EtaA, cyclohexanone for CRE2−CHMO, cyclopentanone for CRE2−CPMO, 4-hydroxy-acetophenone for CRE2−HAPMO and benzaldehyde for CRE2−$PAMO_{M446G}$. Kinetic parameters were obtained by fitting the obtained data as described previously (Kamerbeek *et al.* 2003). Thermostability of the enzymes was determined by incubating approximately 30 µM enzyme at 37 °C or 50 °C, after which their remaining activity was determined at 25 °C using saturating substrate concentrations.

Thermostability measurements of CRE2−PAMO linker variants

Thermostability of the enzymes was determined by incubating approximately 10 µM of enzyme in the presence of 10 µM of FAD at 55 °C, after which their remaining activity was determined at 25 °C. Activity was measured by monitoring the absorbance at 340 nm (NADPH depletion/production) (PAMO: 0,1 µM enzyme, 100 µM NADPH, 2 mM phenylacetone; PTDH: 0,05 µM enzyme, 100 µM $NADP^+$, 5 mM phosphite)

Melting temperature determination of CRE2−PAMO linker variants via ThermoFAD (Forneris *et al*. 2009)

50 µM of purified protein, adjusted with buffer to a volume of 50 µl was placed in a real time PCR (RT-PCR) apparatus (Biorad) equipped with a 470-543 nm excitation filter and a SYBRGreen fluorescence emission filter (523-543 nm). During a temperature gradient from 20 to 90 °C the fluorescence was measured every 0,5 °C, with 10 s delay for stabilization. T_M values are based on the peaks of the derivatives of the fluorscence intensity data.

General protocol for cleared cell extract (CCE) biotransformations

Fresh TB medium (0,2 or 0,5 L) containing 50 µg mL^{-1} ampicillin was inoculated with 1% of an overnight preculture of recombinant *E. coli* TOP10 overexpressing CRE2−PAMO, CRE2−CHMO or CRE2−CPMO a 1 or 2,5 L baffled Erlenmeyer flask, respectively. L-arabinose was added to a final concentration 0.02 % (wv^{-1}) and the culture was incubated at 120 rpm at 24 °C on an orbital shaker for 24 or 40 hours, depending on the total volume. Cells were harvested by centrifugation (6000 × g, 10 min, 4 °C, Sigma 3K30 centrifuge, rotor 19777). The cell pellets were resuspended in 10 mM PBS buffer, pH 7,4 (double volume of cells) containing 0,1 % (vv^{-1}) PMSF and disintegrated by ultrasonication. Cell debris was removed by centrifugation (15000 × g, 20 min, 4 °C, Sigma 6K15 centrifuge, rotor 372/C). The CCE was divided into aliquots of 0,5 mL and stored at –20 °C until needed. Protein concentration of the cleared cell extract was estimated by Bradford method using Protein Assay (Bio-Rad). Biotransformations using CCE were realized in either regular 96-well or 12-well plates. The total reaction volume in each well was 150 µL or 2 mL, respectively, and contained approximately 8 mg mL^{-1} protein and 100 mM Na_2HPO_3 (final concentrations). Substrate inhibition studies were performed using 4-methyl cyclohexanone (5 − 30 mM). Other biotransformations were carried out using 5 or 20 mM racemic bicyclo[3.2.0]hept-2-en-6-one as substrate. The mixture was incubated at 700 rpm at 24 °C in a multi-well plate incubator (Heidolph Titramax 1000). All biotransformations carried out at elevated temperatures (37 °C and 50 °C) were supplemented with 10 µM FAD (final concentration). Samples were taken in regular intervals, extracted with ethylacetate including internal standard (1 mM of methyl-benzoate) and analyzed by gas chromatography (Stewart *et al.* 1996).

Vielleicht wird man eines Tages in den Gebeinen
der Menschen die DNA Gottes entdecken.
Anne Rice

CHAPTER 3

EXPANDING THE SET OF RHODOCOCCAL BAEYER-VILLIGER MONOOXYGENASES BY HIGH-THROUGHPUT CLONING, EXPRESSION AND SUBSTRATE SCREENING

This chapter is based on: ***Applied Microbiology and Biotechnology (2012), 95, 1479-1489***
by A. Riebel, H. M. Dudek, G. de Gonzalo, P. Stepniak, L. Rychlewski and M.W. Fraaije.

Abstract: To expand the available set of Baeyer-Villiger monooxygenases (BVMOs), we have created expression constructs for producing 22 Type I BVMOs that are present in the genome of *Rhodococcus jostii* RHA1. Each BVMO has been probed with a large panel of potential substrates. Except for testing their substrate acceptance, also the enantioselectivity of some selected BVMOs was studied. The results provide insight into the biocatalytic potential of this collection of BVMOs and expand the biocatalytic repertoire known for BVMOs. This study also sheds light on the catalytic capacity of this large set of BVMOs that is present in this specific actinomycete. Furthermore, a comparative sequence analysis revealed a new BVMO-typifying sequence motif. This motif represents a useful tool for effective future genome mining efforts.

3.1. Introduction

In the last few years the interest in using Baeyer-Villiger monooxygenases (BVMOs) as biocatalysts has increased. These flavin dependent enzymes are able to perform efficient chemo-, regio-, and/or enantioselective oxygenation reactions using stoichiometric amounts of O_2 as oxidant and NAD(P)H as electron donor. It has been shown that BVMOs can be used as biocatalysts for highly enantioselective Baeyer-Villiger oxidations and sulfoxidations (de Gonzalo *et al*. 2010). In addition, these oxidative biocatalysts have also been shown to catalyze oxygenation of other heteroatoms like nitrogen, phosphor, selenium and boron, as well as epoxidation reactions (Renz & Meunier 1999, Mihovilovic *et al*. 2004). Such selective oxygenation reactions are difficult to perform selectively with chemical catalysts, rendering BVMOs a powerful tool in organic chemistry.

Taking into account the urgent need for environmental friendly processes, the use of biocatalysts is highly favorable due to the mild reaction conditions possible by employing enzymes such as BVMOs. However, there are still some drawbacks concerning the applicability of these enzymes. One of them is the fact that BVMOs are coenzyme dependent and therefore need an NAD(P)H regeneration system. Several different strategies have been applied to solve this problem, like the use of whole cells or of a regeneration enzyme partner (Hollmann *et al*. 2007, Torres Pazmiño *et al.* 2009). Another complicating factor is the limited availability of suitable BVMOs. With the known BVMOs, only a limited set of reactions can be catalyzed. Furthermore, some of the BVMOs that are of interest are rather unstable. This is already evident from the most studied BVMO: cyclohexanone monooxygenase (CHMO). This BVMO from *Acinetobacter* has been shown to act on a variety of compounds (Stewart 1998, Mihovilovic *et al*. 2002), but it is very unstable (CHMO: $t_{1/2}$ = 24 h at 25°C) (Zambianchi *et al.* 2002). In contrast, phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* is a very robust BVMO ($t_{1/2}$ =24h at 52°C) (Fraaije *et al.* 2005, Secundo *et al*. 2010). Nevertheless, its substrate scope is rather limited to small aromatic substrates.

The above-mentioned current limitations in BVMO-based biocatalysis has triggered enzyme discovery and engineering studies to generate a larger collection of BVMOs that

show a good operational stability, while covering a wide variety of compounds. One approach to achieve this is to engineer known BVMOs in order to create supplementing activities and selectivities. For this purpose, several enzyme engineering studies have been reported in recent years (Fraaije *et al*. 2005, Torres Pazmiño *et al*. 2010). For example, PAMO has been subjected to structure-based mutagenesis and this has resulted in several mutants that display a somewhat altered substrate range and/or enantioselectivity when compared with the wild-type enzyme (Bocola *et al.* 2005, Torres Pazmiño *et al.* 2007, Wu *et al.* 2010, Dudek *et al*. 2011). However, these enzyme engineering studies have also shown that it is difficult to introduce substrate acceptance profiles that are significantly different from the parent enzyme.

To increase the number of available and distinct BVMOs, also the natural diversity can be tapped. With the BVMO-specific sequence motif it has become easy to survey sequenced genomes for BVMO-encoding genes (Fraaije *et al.* 2002). Such identified genes can subsequently be cloned and the corresponding proteins overexpressed for biocatalytic evaluation studies. Genome mining has already delivered a number of novel BVMOs (Torres Pazmiño *et al*. 2010). Even though BVMOs are widespread in bacteria and fungi, the average distribution is only around 1 BVMO-encoding gene per microbial genome (de Gonzalo *et al.* 2010). In this respect the bacterium *Rhodococcus jostii* RHA1 is a peculiar and promising exception: its predicted proteome contains an unusually high variety in oxidative enzymes, including >20 putative BVMOs (McLeod *et al.* 2006). The Grogan group already performed an initial study on the BVMO repertoire of this organism (Szolkowy *et al.* 2009) and identified several interesting activities but failed in expressing all BVMOs.

In this chapter, we report on the exploration of 22 BVMOs derived from R. jostii RHA1. By optimizing expression vectors and expression conditions all BVMOs were obtained in a soluble form. Using a newly developed method to determine the amount of expressed and active BVMO in a cell extract in combination with a novel

generic activity screening assay, the substrate profiles for a large set of substrates for all 22 BVMOs have been determined.

3.2. Results

3.2.1. Identification of putative BVMO-encoding genes

By using protein sequences of PAMO and CHMO as baits for BLASTP searches in the predicted proteome of *R. jostii* RHA1 (McLeod *et al.* 2006), and inspecting individual putative sequences for known sequence motifs, 23 putative Type I BVMO genes could be identified (Table 3-1). For consistency, we numbered the BVMOs according to the numbering employed in a previous study by the Grogan group (Szolkowy *et al.* 2009). Compared with this study, our set of BVMOs contains one more putative BVMO-encoding gene that was not identified before (BVMO #24). Furthermore, it was previously suggested that BVMO #8 represented a BVMO lacking a large part of the C-terminal domain, which binds the NADPH coenzyme (RHA1_ro08998). Careful inspection of the sequenced genome revealed that a large part of a C-terminal BVMO sequence (ORF RHA1_ro08999) resides upstream of RHA1_ro08998 in the genome. Sequence analysis revealed that only one nucleotide was missing to create a gene encoding a full length BVMO, fusing both these open reading frames. The corrected BVMO #8 gene was generated and included in this study (see Materials and Methods). Except for these two newly reported BVMOs, BVMO #8 and BVMO #24, we also discarded one gene from our study that was previously assigned as a Type I BVMO (BVMO #22 in Szolkowy *et al.* 2009). Analysis of the respective protein sequence revealed that it represents a typical flavin-containing monooxygenase (FMO), which is a member of a flavoprotein family that is distinct from BVMOs (Fraaije *et al*. 2002). Furthermore, two BVMOs (BVMO #21 and BVMO #23) differed in only one amino acid out 546 amino acids. Based on sequence alignment with phenylacetone monooxygenases, this residue (Leu or Arg at position 480) is predicted to be on the surface of the protein and therefore will have little effect on the catalytic properties of these BVMOs. This has in fact been confirmed by the previous study

of the Grogan group (Szolkowy *et al*. 2009). For that reason, BVMO #23 was excluded from this study.

All identified BVMO sequences contained two conserved Rossmann fold motifs: one near the N-terminus and one in the middle of the sequence (Table 3-1). In BVMO #10 and BVMO #18, the N-terminal GxGxxG motif is preceded by ~120 amino acids instead of only ≈20 amino acids. This indicates that these enzymes have acquired an additional domain that is also found in a known BVMO, 4-hydroxyacetophenone monooxygenases (HAPMO), and which has been suggested to play a role in dimerization (Kamerbeek *et al*. 2005). As described before, Type I BVMOs contain a typifying short sequence motif: FxGxxxHxxxW[P/D] (Fraaije e*t al.* 2002). By allowing one or two conservative mutations in this motif, all 22 sequences were found to contain such sequence motif. It shows that not every residue of the previously described motif is strictly conserved. While the glycine and tryptophan residues are strictly conserved, some mutations (*e.g.* F>Y in BVMO #4), also occur which reflect the origin of the sequence conservation within this motif. The BVMO motif is located in a loop region of BVMOs that allows protein dynamics during the catalytic cycle while it is not directly involved in catalysis (Orru *et al.* 2011).

For illustrating the sequence relationships between the identified *R. jostii* sp. RHA1 BVMOs and some of the best characterized BVMOs, all these sequences were aligned (see Appendix). This revealed that the rhodococcal BVMOs are representatives of many different clades of the Type I BVMO family. Careful inspection of the structure-based multiple sequence alignment also revealed another interesting conserved sequence motif. Except for the two Rossmann fold motifs and the BVMO-motif mentioned above, there is one particular region of the sequences that shows remarkable conservation, located between the N-terminal GxGxxG motif and the known BVMO-motif. This region can be defined by [A/G]GxWxxxx[F/Y]P[G/M]xxxD. The full conservation of the aspartate is in full agreement with the recent finding that Asp66 in PAMO plays a crucial role in catalysis by direct interaction with the $NADP^+$ coenzyme. Mutating this residue into an alanine resulted in inactivation of the enzyme by decreasing the rate of NADPH-mediated flavin reduction (Orru *et al.* 2011). Combined with Arg337, these two residues represent the

only fully conserved active site residues, both crucial for catalysis. The conserved doublet of glycine residues in this newly recognized motif is a known motif for a Rossmann fold domain and in the BVMOs they support effective binding of the FAD cofactor. The other conserved residues in the motif center around the flavin cofactor and appear to be essential for creating a productive active site architecture (Figure 3-1).

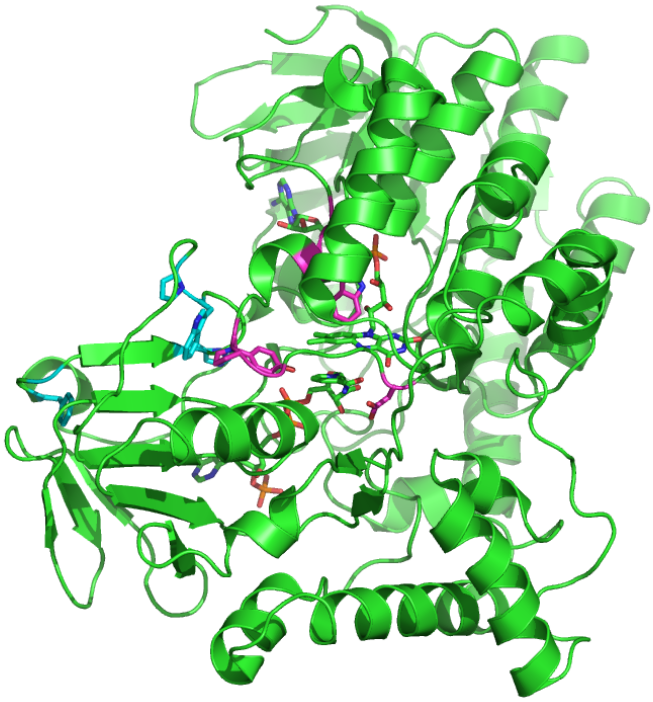


Figure 3-1: Structure of PAMO with conserved residues highlighted. The previously described BVMO motif residues are in blue sticks while the motif described in this chapter is in magenta sticks. The FAD and $NADP^+$ cofactors are shown in green sticks.

This newly recognized motif appears to be more discriminative for identifying BVMO sequences as it entails more and more conserved residues. It also clearly differentiates between BVMO and FMO sequences as the structure and sequence region around Asp66 (in PAMO) in FMOs is totally different (Fraaije *et al*. 2002, Alfieri *et al*. 2008). Therefore, this motif represents a new and more reliable tool to annotate putative BVMO-encoding genes.

In order to obtain a better view on the sequence-function relationship among the tested set of BVMOs, also another sequence alignment was performed. Using the recently elucidated crystal structure of PAMO, which contains $NADP^+$ and a substrate analogue (MES), we identified all residues that are within a 12 Å radius from the sulphur atom of the substrate analogue (Orru *et al.* 2011). This yielded a set of 59 residues that form the core of the substrate binding pocket and that are spread over the PAMO sequence. Upon performing a structure-based multiple sequence alignment, we filtered for all respective 59 residues for generating a separate multiple sequence alignment. We anticipated that such an approach of comparing partial sequences would have a higher probability to disclose similarities in substrate specificities. In Figure 3-2, the phylogenetic tree representation of this focussed sequence alignment of all tested and known BVMOs is shown. The grouping differed slightly from the global multiple sequence alignment and was used for the discussion below.

3.2.1. Expression of the putative BVMOs

All 22 putative BVMO-encoding genes were cloned into a slightly modified pBAD expression vector (Kamerbeek *et al.* 2004) using a ligation free cloning method. Subsequently, expression tests were performed in which the temperature and arabinose concentration were varied. With these constructs, 11 BVMOs could be obtained in soluble form with good expression levels, using *E. coli* TOP10 as host (Table 3-1). A similar success rate was previously obtained for roughly the same set of genes using an pET-based *E. coli* expression system (Szolkowy *et al.* 2009) (Table 3-1). In the other cases, the expressed protein was either insoluble or the expression levels of soluble protein were very low. To increase the expression level and/or solubility, the 11 BVMO genes that could not be expressed using pBAD were cloned into the recently developed pCRE2 expression vector and tested using various expression conditions (Torres Pazmiño *et al.* 2009). This resulted in good expression levels of soluble protein of most of these genes. Only in 5 cases poor soluble protein expression was observed. Nonetheless, by using the NADPH-dependent flavin reduction method (*vide infra*) it was found that also these constructs led to

expression of detectable levels of soluble NADPH-reactive proteins. Taken together, by using two expression vectors and careful tuning of expression conditions, we succeeded in producing all 22 targeted BVMOs in a soluble form.

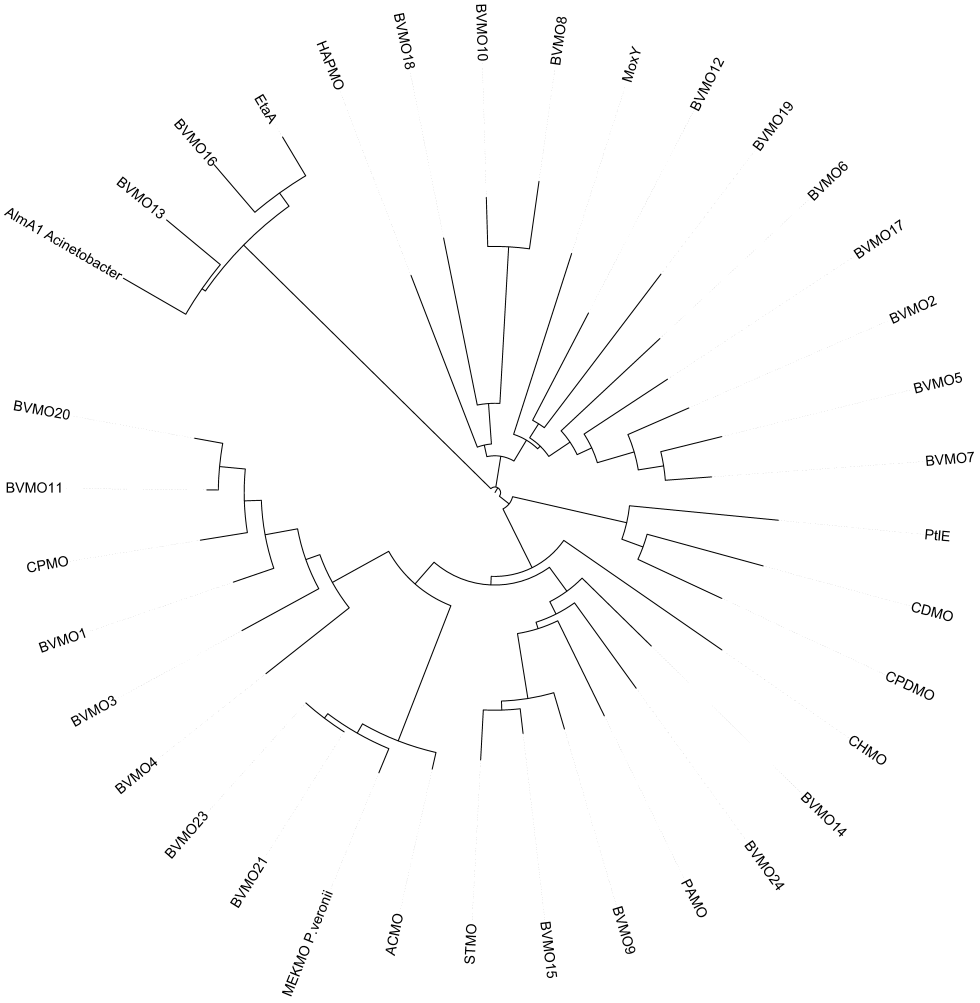


Figure 3-2: Radial phylogenetic representation of the comparison of active site residues of the selected BVMO protein sequences. The included sequences are: AlmA1, *Acinetobacter* sp. DSM 17874 (ABQ18224.1); EtaA, *Mycobacterium tuberculosis* H37Rv (NP_218371.1); HAPMO, *Pseudomonas fluorescens* ACB (Q93TJ5.1); MoxY, *Aspergillus flavus* (AAS90037.1); PtlE, *Streptomyces avermitilis* MA-4680 (NP_824170.1); CDMO, *Rhodococcus ruber* SCI (AAL14233.1); CPDMO, *Pseudomonas* sp. strain HI-70 (BAE93346.1); CHMO, *Acinetobacter calcoaceticus* NCIMB 9871 (BAA86293.1); PAMO, *T. fusca* (YP_289549.1); STMO, *Rhodococcus rhodochrous* (BAA24454.1); ACMO, *Gordania* sp. strain TY-5 (BAF43791.1); MEKMO, *Pseudomonas veronii* MEK700 (ABI15711.1); CPMO, *Comamonas* sp. strain NCIMB 9872 (Q8GAW0).

Table 3-1: All identified BVMOs and their sequence characteristics from the predicted *R. jostii* RHA1 proteome.

BVMO	*accession number*	*Rossman motif* GxGxxG	*BVMO motif* FxGxxxHxxxW$^{P}/_{D}$	*Rossman Motif* GxGxxG	*length (aa)*	*soluble* [b] *expression*	*medium color*
1	ro06679	GxGxxG	FxGxxxHxxxWP	GxGxxG	538	+ B	-
2	ro04304	GxGxxG	FxGxxxHxxxWN	GxGxxA	496	++ C	-
3	ro03247	GxGxxG	FxGxxxHxxxWP	GxGxxG	543	++ C	blue
4	ro03063	GxGxxG	YxGxxxHxxxWR	GxGxxA	549	++ B	-
5	ro02109	GxGxxG	FxGxxxHxxxWN	GxGxxA	491	+ C	blue
6	ro01874	GxGxxG	FxGxxxHxxxWN	GxGxxA	493	++ B	-
7	ro06008	GxGxxG	FxGxxxHxxxWN	GxGxxA	537	++ C	blue
8[a]	ro08998/9	GxGxxG	FxGxxxHxxxWD	GxGxxA	545	++ C	-
9	ro09035	GxGxxG	FxGxxxHxxxWP	GxGxxG	541	++ B	-
10	ro09039	GxGxxG	FxGxxxHxxxWD	GxGxxS	663	++ B	-
11	ro06698	GxGxxG	FxGxxxHxxxWP	GxGxxG	545	++ B	-
12	ro07112	GxGxxG	FxGxxxHxxxWD	GxGxxA	514	++ C	blue
13	ro03773	GxGxxG	FxGxxxHxxxWP	GxGxxA	515	+ C	-
14	ro03437	GxGxxG	FxGxxxLxxxWP	GxGxxG	547	++ B	-
15	ro02492	GxGxxG	FxGxxxHxxxWP	GxGxxG	553	++ B	-
16	ro02919	GxGxxG	FxGxxxHxxxWP	GxGxxG	498	++ C	-
17	ro05228	GxGxxG	FxGxxxHxxxWD	GxGxxA	603	+ C	blue
18	ro05396	GxGxxG	FxGxxxHxxxWP	GxGxxS	652	++ C	-
19	ro05522	GxGxxG	FxGxxxHxxxWN	GxGxxA	515	+ C	brown/blue
20	ro08137	GxGxxG	FxGxxxHxxxWP	GxGxxG	548	++ B	brown/blue
21	ro10187	GxGxxG	FxGxxxHxxxWP	GxGxxG	546	++ B	brown/blue
23	ro08185	GxGxxG	FxGxxxHxxxWP	GxGxxG	546	n.d	n.d.
24	ro05323	GxGxxG	FxGxxxSxxxWN	GxGxxG	564	++ B	yellow

[a], this gene has been cured, see text; [b], ++ indicates a clear visible protein band using SDS-PAGE while + indicates that no clear band was visible but flavin reduction was observed, B indicates that the respective BVMO has been expressed using the pBADN-vector, C indicates the use of the pCRE2- expression vector; the underlined BVMOs were previously expressed in soluble form in a pET expression vector (Szolkowy *et al.* 2009). n.d.: not determined.

Interestingly, for a number of constructs it was found that, using the optimal expression conditions, the growth medium changed significantly in color (Table 3-1). Such a phenomenon had been observed before in a BVMO mutagenesis study where the M446G PAMO mutant was created (Torres Pazmiño *et al.* 2007). This mutant BVMO was found to form indigo blue by oxidizing indole, which is present in the cytosol of *E. coli*. Also for some of the expressed rhodococcal BVMOs a blue color was observed while some other constructs led to a more brownish or intense yellow color. Apparently, some of the

expressed BVMOs act on endogenous compounds present in the culture medium or in the cells, which confirms functional expression of these monooxygenases.

3.2.2. Substrate profiling

To establish the biocatalytic potential of this large set of BVMO, we aimed at testing each of these BVMOs for activity on a large panel of potential substrates. For such substrate profiling study, we first set out to quantify the amount of active enzyme in cell extracts. Based on the fact that all Type I BVMOs contain a FAD cofactor that is specifically reduced by NADPH, we developed a simple method for accurate BVMO concentration determination. This approach is somewhat analogous to the well-known method to quantify the amount of P450 monooxygenases by using the specific heme absorption features of reduced P450 monooxygenases with carbon monooxide. In *E. coli* only a minority of the native proteins contain a flavin cofactor and most of them will not be reactive with NADPH. In contrast, the overexpressed BVMOs will be dominant in the respective extracts, outnumbering the endogenous flavoproteins of *E. coli*. By measuring the absorbance spectrum of the CCE before and after the addition of NADPH, the amount of NADPH-reactive flavin cofactor in the expressed BVMO can be determined (Figure 3-3).

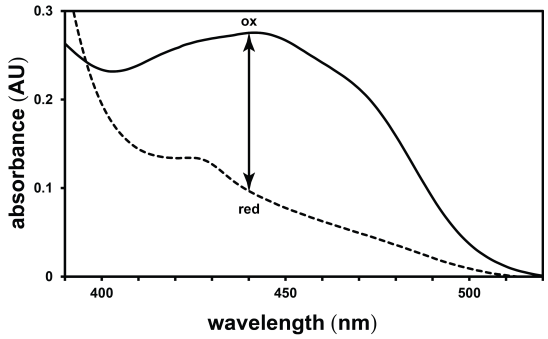


Figure 3-3: Absorbance spectra of a cleared cell extract containing overexpressed BVMO. The spectra of the oxidised flavin cofactor in BVMO #14 (—) and after reduction by NADPH (▪▪▪▪) are shown. The difference in absorbance at 440nm that can be used to determine the BVMO concentration is indicated.

For the calculation an extinction coefficient is needed which reflects the absorbance difference between an oxidized and reduced BVMO. For this, we used the extinction

coefficients for oxidized PAMO ($\varepsilon_{441nm\ ox}$ = 12,3 $mM^{-1}cm^{-1}$) and reduced PAMO ($\varepsilon_{441nm\ red}$ = 3,6 $mM^{-1}cm^{-1}$) to calculate an estimate of the difference extinction coefficient for FAD containing BVMOs at 440 nm ($\varepsilon_{440nm,\ BVMO\ diff.}$ = 8,8 $mM^{-1}cm^{-1}$). We decided to use the absorption at 440 nm and to neglect the small differences caused by different flavin microenvironments in individual BVMOs. This corroborates with flavin absorbance spectral data of known BVMOs from literature, which display highly similar absorbance maxima and extinction coefficients (Table 3-2).

Table 3-2: Collection of literature values for λ_{max} and ε_{ox} of known BVMOs.

BVMO	***PAMO***	***HAPMO***	***CHMO***	***STMO***	***CPMO*** [a]	***CPDMO*** [a]
literature	*Fraaije et al. 2005*	*Kamerbeek et al. 2001*	*Donoghue et al. 1976*	*Miyamoto et al. 1995*	*Griffin et al. 1976*	*Iwaki et al. 2006*
λ_{max} (nm)	**441**	**439**	**440**	**438**	**444**	**442**
ε_{ox} ($mM^{-1}cm^{-1}$)	**12,3**	**12,4**	**13,8**	**11,3**	**14,5**	**12,8**

[a] **these BVMOs were purified as PTDH fusion enzymes and analyzed spectrophotometrically in this study.**

Validation of the method by using purified enzyme with a known concentration showed that the method is accurate. Furthermore, using extracts of *E. coli* devoid of any expressed BVMO did not show any significant A_{440} change after incubation with NADPH, indicating that there is no interfering absorbance. Using this method, it was possible to determine the concentration of soluble BVMO in all 22 samples, independent of the level of BVMO overexpression. In this way, we could normalize the amount of recombinant enzyme in the substrate profiling experiments while using cleared cell extracts.

For substrate profiling we used 39 substrates, representing different compound classes: linear aliphatic ketones, cyclic aliphatic ketones, aromatic ketones, aromatic amines, and aromatic sulfides (Table 3-3). All substrates were used at two different concentrations (5 and 1 mM) and a fixed amount of cosolvent was used. For determining activity for each tested compound, we initially used the NADPH depletion assay. Consumption of NADPH by BVMO can be monitored by measuring the absorbance of NADPH at 340 nm. However, this assay could not be used for compounds that absorbed in the same wavelength regime

of NADPH (300-390 nm) and only allowed to measure the conversion of relatively small concentrations of NADPH and substrate. Therefore, we decided to use an assay that is coupled to the activity of the employed regeneration enzyme, which allows to monitor conversion over a longer time period. For this, phosphite dehydrogenase was used which generates one molecule of phosphate for the regeneration of each NADPH molecule. The formed phosphate could be quantified by using a chromogenic reaction, allowing spectrophotometric detection of the degree of conversion. The conversions were incubated for two hours before measuring the phosphate concentrations. As BVMOs also can act as NADPH oxidases in the absence of a suitable substrate, formation of phosphate was only regarded as significant when it was significantly more when compared with the incubation without a test compound. By using this approach, poor substrates will be overlooked while establishing profiles of efficient substrates for all tested BVMOs. The observed substrate profiles for all tested BVMOs are summarized in Table 3-3.

The results disclose a broad substrate scope for the produced set of 22 BVMOs. To verify the results, we checked a selection of substrates for conversion with several BVMOs by GC analysis, which confirmed the majority of the identified substrates. For most of the tested compounds, we could find at least one active enzyme, even though the intensity of the activity varies a lot between the tested BVMOs. Nonetheless, for 8 BVMOs (BVMO #1, #3, #6, #7, #8, #13, #16, and #17) no significant activity was observed for any of the tested compounds. However, it is worth noting that of these 8 BVMOs, BVMO #3, BVMO #7, and BVMO #17 appear to produce some indigo blue when expressed in *E. coli* (Table 3-1). This suggests that these BVMOs are indeed active while not very efficient on any of the compounds in our test panel. Another relevant observation is the fact that conversions using BVMO #3, BVMO #8, BVMO #13 and BVMO #19 in the absence of substrate results in a 4-5 fold higher background reaction when compared with all other tested BVMOs (≈9 mM phosphate formed instead of ≈2 mM). For BVMO #13 and BVMO #19 this can be explained by the fact that their expression was poor, which dictated the use of relatively high amounts of extract, which leads to a relatively high background NADPH consumption. Contrarily, BVMO #3 and BVMO #8 are well expressed. This suggests that these BVMOs

display an exceptionally high uncoupling rate and therefore substrates may be unnoticed by monitoring NADPH consumption or phosphate formation. It also agrees with the fact that Szolkowy *et al.* (2009) have found that BVMO #3 is indeed active on a number of ketones. For the BVMOs that did not display significant background NADPH consumption or uncoupling rates, the substrate acceptance profiles varied considerably. Among the tested BVMOs, only a small set of 5 BVMOs were shown to be active on ≥10 tested compounds, leaving only 5 potent BVMOs respectively. This contrasts the catalytic potential of well characterized BVMOs.

As known from literature, CHMO, CPMO, PAMO and HAPMO were indeed found to accept a broad range of compounds. Only STMO, for which a substrate profile had not yet been established, appeared rather selective by converting only 9 of the tested compounds. Similar to the known prototype BVMOs, two rhodococcal BVMOs, BVMO #4 and BVMO #24, appeared to be very substrate tolerant as they both accepted the majority of the tested compounds. Their substrate profiles resemble to a large extend that of CPMO, with BVMO #24 exhibiting greater efficiency when compared with BVMO #4. It is worth noting that BVMO #24 had escaped the previous rhodococcal BVMO exploration study by the Grogan group. Of the other BVMOs, BVMO #2 and BVMO #20 also were found to be active on a number of cycloalkanones. BVMO #9 and BVMO #15 showed high activity with a number of aromatic ketones, resembling to some extend the substrate scope of HAPMO. Other BVMOs appeared to be rather restricted in substrate scope. For example, BVMO #10 appears to be very selective in acting on bulky ketones converting bicycloheptenone and phenendione. Still, these BVMOs may be valuable biocatalysts exhibiting new substrate specificities and chemo-, regio-, or enantioselectivitiesthat were also tested with the panel of test compounds (Table 3-3).

Table 3-3: Substrate profiling. Activities as measured by phosphate formation are indicated for the rhodococcal BVMOs and prototype BVMOs. BVMOs that did not show any activity on the tested compounds were excluded. The rhodococcal BVMOs are indicated by their respective numbers. The observed activities are indicated as +, ++ or +++ and reflect 1,2, 2 or 5 fold the phosphate formation (cq. conversion) when compared with incubations lacking the tested compound.

Substrates[b]		rhodococcus BVMOs														standard BVMOs[a]				
		2	4	5	9	10	11	12	14	15	18	19	20	21	24	PA	CP	CH	ST	HAP
1	Acetone		+								+				++		+			+
2	2-Butanone		+												++		+			+
3	2-Octanone		+	+		+							+	+	++		++	+		+++
4	2-Dodecanone		++	+									+				+++	+++		+++
5	3-Octanone		+												++		++			+++
6	4-Octanone	+	+									+			++		++	+		+
7	Methylvinylketone		+		+					+			+	+	++	+	++		+	+++
8	3-Methyl-2,4-pentanedione		+							+	+		+		++		++			+++
9	Cyclobutanone		+							+			+		++		++			+++
10	Cyclopentanone	+	+										+	+	++		++	++		+
11	Cyclohexanone	+	+										+		++		+++	++		+
12	Cyclopentadecanone		+														++	+		+
13	Oxocyclohexanecarbonitrile	+	+												++		++			+++
14	4-Methylcyclohexanone		+														+++	++		+
15	2-Propylcyclohexanone		+	+											+		+++	++	+	+
16	Dihydrocarvone		+			+									+		++	++		+++
17	Cyclopropylmethylketone	+	+										+	+			++			+++
18	Norcamphor		+										+	+	++		+++	++		+
19	Bicycloheptenone					+++		+					+	+	++		+++	++		+++
20	Thioanisole		+								+						++	++		+++
21	Benzylethyl sulfide		+						+	+					++		++	++		+++
22	Benzylphenyl sulfide		+						+						++		++	+		+++
23	Ethionamide		+		++					+					++	+++	+		+++	+++
24	Modafinil precursor		+												+	+++	+			
25	Thiacetazone												+	+	++	+	+			+
26	Nicotin		+				+										++	+		
27	Indole															+	+			
28	3-Acetylindole				++					+						+		+		+++
29	Acetophenone		+										+		++		++		+++	+++
30	4-Hydroxyacetophenone		+		+	+				+++					+	+	++			+++
31	2,6-Dihydroacetophenone		+													+	++		+++	
32	3-Phenylpentane-2,4-dione									+					+	+++	++		+	+
33	Phenylacetone		+		+					+++			+	+		+++	++		+	+++
34	Raspberry ketone		+		++					+				+	+	+++	+	+	+++	
35	2-Phenylcyclohexanone																+	++	++	+
36	Benzoin																+			+++
37	Phenendione					+++										+	+			
38	2-Indanone	+													+		++	+++		+
39	1-Indanone												+	+			++			+
in total converted substrates		6	29	3	6	5	1	1	2	10	3	1	14	11	25	12	38	19	9	33

[a], The abbreviations used are PA: PAMO, CP: CPMO, CH: CHMO, ST: STMO, HAP: HAPMO; [b] The structural formulas and other details can be found in Appendix A-2. The shading indicates different substrate classes.

.

To have a better insight into the biocatalytic potential of some of the BVMOs, a selected set was studied for the conversion of some typical BVMO substrates, including prochiral sulfides. Table 3-4 shows that all these BVMOs convert the typical BVMO substrates phenylacetone and bicycloheptenone.

Table 3-4: GC analysis of substrate conversions by selected BVMOs.

BVMO	*3*	*8*	*9*	*14*	*15*	*20*	*21*	*24*
phenylacetone	**99**	**27**	**99**	**99**	**92**	**45**	**99**	**38**
2-indanone	**<3**	**<3**	**<3**	**<3**	**<3**	**<3**	**<3**	**62**
bicycloheptenone	**97**	**<3**	**36**	**97**	**90**	**97**	**97**	**93**
thioanisole	**87**	**<3**	**21**	**99**	**76**	**24**	**55**	**80**
enantioselectivity	**90 (S)**	**n.d.**	**61 (S)**	**n.d.**	**45 (S)**	**5 (S)**	**63 (S)**	**82 (R)**
benzylethyl sulfide	**60**	**<3**	**<3**	**72**	**14**	**4**	**92**	**91**
enantioselectivity	**75 (R)**	**n.d.**	**n.d.**	**90 (R)**	**20 (R)**	**n.d.**	**93 (R)**	**>99 (S)**

BVMO #8 gives poor conversion of phenylacetone while none of the other compounds are converted. This confirms the outcome of the activity screening, which did not reveal any substrate for this enzyme. That the conversion of phenylacetone can be observed is due to the fact that the reactions have been performed over a relatively long time period. For 2-indanone, only BVMO #24 was found to be reactive which is also in line with the substrate profiling study (Table 3-3). Except for BVMO #15, BVMO #20 and BVMO #8, the two aromatic sulfides were efficiently converted with good to excellent enantio-selectivities. Strikingly, most of the BVMOs show the same enantioselectivity forming predominantly one specific sulfoxide enantiomer of each of the two aromatic sulfides. Only BVMO #24 is again an exception and shows opposite enantioselectivity.

3.3. Discussion

By employing two expression vectors and screening for proper expression conditions, we have been able to express 22 Type 1 BVMOs from *R. jostii* RHA1. This is the largest set of recombinant BVMOs produced from one microbial origin reported so far. The collection supplements the previous set of 13 BVMOs produced by the Grogan group. Except for establishing expression conditions for BVMOs that could not be expressed before, we also

identified two new BVMO genes (BVMO #8 and BVMO #24). The biocatalytic exploration of this new collection of BVMOs was aided by the development of two new methods: (*i*) by exploiting the distinct absorbance features of oxidized and reduced protein-bound FAD, the BVMO content in extract could be determined; and (*ii*) by performing the conversions coupled to NADPH regeneration by phosphite dehydrogenase, conversion could be quantified by colorimetric phosphate detection. Sequence analysis also revealed a novel BVMO-typifying motif: [A/G]GxWxxxx[F/Y]P[G/M]xxxD. This motif includes residues that are part of the active site and therefore represents a useful tool for annotating BVMO-encoding genes. The motif is absent in FMO sequences and therefore allows discrimination between members of the two sequence-related flavoprotein subfamilies, BVMOs and FMOs.

The unraveled substrate profiles of the studied BVMOs differed significantly among the produced BVMOs. For some BVMOs no substrate could be identified using a test set of 39 compounds. This could indicate that the natural substrates for these enzymes differ too much from the tested compounds. However, it may also be that these BVMOs only display very low oxygenation activities that could not be detected with the employed assay. Such low activity has also been observed for EtaA from *Mycobacterium tuberculosis* (Fraaije 2003). This drug-resistance related enzyme was shown to act on a variety of typical BVMO substrates but with very low activity. A small number of the produced BVMOs were active on ≥10 compounds: BVMO #4, BVMO #15, BVMO #20, BVMO #21, BVMO #24. BVMO #24 appeared to be a very potent biocatalyst as it showed very good activities on 26 compounds. BVMO #3 may also belong to this group of potent BVMOs but appears to act as NADPH oxidase which prevented establishing its true substrate profile. To correlate sequence to substrate profiles, we generated a phylogenetic tree that is based on the comparison of residues that make up the core of the active site. This revealed that all BVMOs that displayed a relaxed substrate acceptance profile were closely related to the clades that include the well-known prototype BVMOs (CHMO, CPMO, PAMO and STMO). The above-mentioned potent rhodococcal BVMOs, including BVMO #3, are all closely related to these prototype BVMOs (Figure 3-2). The rhodococcal

BVMOs that are part of sequence related groups that include other known but less explored BVMOs (HAPMO, MoxY, EtaA, AlmA) appear to be less promising biocatalysts. This observation is valuable for future BVMO discovery efforts: it appears more productive to search for homologs of the typical prototype BVMOs when looking for potent biocatalysts.

Two BVMOs display a peculiar behavior: BVMO #3 and BVMO #8 act as NADPH oxidases when no potential substrate is present. For the other BVMOs and all previously characterized Type 1 BVMOs, this futile oxidation of NADPH was typically very slow. The relatively high NADPH oxidase activity may reflect the fact that these respective BVMOs do not function properly when taken out of their natural context. Alternatively, it may indicate that these enzymes have acquired NADPH oxidase activity for a certain purpose, *e.g.* to regenerate $NADP^+$ in the cell or to produce H_2O_2. Production of H_2O_2 complies with the recent discovery that actinomycetes are relatively rich in peroxidases (van Bloois *et al.* 2010). In line with this, a DyP-type lignin peroxidase from *R. jostii* sp. RHA1 has recently been identified and characterized (Ahmad *et al.* 2011). Except for displaying NADPH oxidase activity, BVMO #3 also was found to be acting as an oxygenase as the cells expressing this enzyme developed a blue color indicative of indole oxygenation. This would imply that BVMO #3 and BVMO #8 have a dual function in *R. jostii* RHA1: acting as oxygenases or NADPH oxidases. In this context, it is worth noting that, due to a missing nucleotide, the gene encoding BVMO #8 is disrupted. This may reflect that due to the relative high NADPH oxidase activity, generating toxic hydrogen peroxide, it was beneficial to silence the gene. Further work on the enzymes reported in this study will disclose details on their catalytic properties and biocatalytic potential.

3.4. Experimental Section

Reagents & enzymes

Oligonucleotide primers were purchased from Sigma, dNTPs and In-Fusion™ 2.0 CF Dry-Down PCR Cloning Kit from Clontech, Phusion polymerase from Finnzymes. All other chemicals were obtained from Acros Organics, ABCR, Sigma-Aldrich, TCI Europe, and Roche Diagnostics GmbH.

Bacterial strains and Plasmids

Escherichia coli TOP10 from Invitrogen was used as a host for DNA manipulations and protein expression. Two expression vectors have been used: (*i*) a modified pBAD vector (pBAD*N*) in which the NdeI site was replaced by the original NcoI site (Kamerbeek *et al*. 2004), and (*ii*) the recently engineered pCRE2 vector which harbors a codon-optimized gene encoding a 18x mutant of phosphite dehydrogenase (PTDH) with an N-terminal His-tag (Torres Pazmiño *et al*. 2009).

Sequence analysis and cloning

The NCBI server was used for DNA sequence retrieval and BLAST searches (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The preliminary alignment of BVMOs was prepared using ClustalW multiple alignment software (Thompson *et al.* 1994). Next, secondary structure information was extracted from the recently elucidated PAMO structure containing $NADP^+$ and a bound substrate analogue (PDB: 2YLT, chain A) *(Orru et al.* 2011) and used for adjustments of the alignment. For calculating the Neighbor-Joining trees and bootstrapping (1500 iterations), we applied the MEGA5 software package (Tamura *et al.* 2011). The phylogenetic trees were visualized by using the iToL online tool (Letunic & Bork 2011).

Target genes were amplified using the genomic DNA of *R. jostii* RHA1 (originally isolated by Seto et al in 1995 (Seto *et al*. 1995)) as template. The genome of *R. jostii* RHA1 has been fully sequenced (McLeod *et al*. 2006) and information can be found at www.rhodococcus.ca. The genes were cloned into pBAD or pCRE2 expression vectors

using the In-Fusion PCR Cloning kit from Clontech, following the recommendations of the manufacturer. Gene #8 was cured by inserting a nucleotide (C) at position 621. In that way the stop codon of the ORF RHA1_ro08999, encoding a putative N-terminal domain of a Type I BVMO, was changed into a codon for leucine. This amino acid was found most frequently in BVMOs at the respective position. By the subsequent frame shift ORF RHA1_ro08999 and RHA1_ro08998 are fused at DNA level encoding a full length Type I BVMO. The insertion was achieved by QuikChange® site-directed mutagenesis using the following primers: forward 5'-caccggtttcgggttcctgaactcccccaa-3'
reverse 5'-ttgggggagttcaggaacccgaaaccggtg-3'

Gene expression

Expression was tested using 24 deep square well microtiter plates in the sandwich cover system from Enzyscreen®, shaking at 200 rpm. Cell cultures of 2,5 ml in Luria–Bertani medium supplemented with 50 µg mL^{-1} ampicillin were grown at four different temperatures (17°C for 48 h, 24°C for 32 h, 30°C and 37°C for 16 h) with three different inducer (arabinose) concentrations (0,002%, 0,02%, 0,2%, 0% as control) each. Cell extracts were obtained using DNase/lysozyme in combination with freezing in liquid nitrogen and thawing at 30°C. Cell extracts (CE) as well as cleared cell extracts (CCE) were analyzed via SDS-PAGE for (soluble) expression of the respective BVMOs.

BVMO concentration determination in cleared cell extract

For the substrate screening assays, cell cultures of 500 – 800 mL Luria–Bertani medium supplemented with 50 µg mL^{-1} ampicillin and the corresponding amount of arabinose were grown. Cell extract (CE) was prepared using sonication with subsequent centrifugation to remove the cell debris and obtain a cleared cell extract (CCE). An absorption spectrum of the CCE was measured from 700 nm to 300 nm (oxidized FAD), after which 500 µM NADPH was added. The sample was incubated at 25°C for 10 min for full flavin reduction after which again a spectrum was measured. Subtracting the two collected spectra revealed the amount of NADPH-reactive flavin in the CCE. To minimize interfering absorbance due to protein aggregates of *E. coli* proteins, the absorption at

440 nm was corrected by subtracting the absorption value at 700 nm. The amount of flavin-containing BVMO in the CCE was calculated using the difference extinction coefficient derived from PAMO: $\varepsilon_{difference}$= 8,8 $mM^{-1}cm^{-1}$.

Substrate screening

For activity screening, an indirect activity assay was used. The substrates were dissolved in DMSO or dioxane. The solvent (DMSO or dioxane) concentration was kept at 5% in all tested reactions. For the activity assay, 100 µL 50 mM Tris/HCl pH 7,5, containing 1 or 5 mM substrate, 100 µM NADPH, 10 µM PTDH, 10 mM sodium phosphite and 1 µM BVMO as CCE was incubated at 25°C for two hours. 20 µL of the reaction mixture was subsequently mixed with 200 µL molybdate reagent (10 mM $(NH_4)_6Mo_7O_{24}$ x 4 H_2O, 100 mM Zn acetate, pH 5, adjusted with HCl) and 50 µl 10% ascorbic acid (pH 5, adjusted with 40% NaOH). During incubation of 30 min at 30°C a blue colored complex formed due to the phosphate generated during the oxidation reaction (Saheki *et al.* 1985). By measuring the absorption at 600 nm, the concentration of produced phosphate could be quantitatively determined using a calibration curve. By being able to quantify the amount of applied BVMO each conversion could be performed with a standard amount of BVMO which allows a direct comparison.

GC analysis

For GC analysis, 500 µL incubations of 50 mM Tris/HCl pH 7,5, containing 5 mM substrate, 5% cosolvent (DMSO or dioxane), 100 µM NADPH, 3,1 µM PTDH, 10 mM phosphite and 5 µM BVMO (CCE) was incubated shaking at 24°C for 2 hours. The reactions were then stopped by extracting with ethyl acetate (2 x 0,5 mL including 0,1% mesitylene as an internal standard), dried with magnesium sulfate and analyzed directly by GC to determine the degree of conversion. The details concerning the (chiral) GC analysis can be found in the Appendix A-1. For every tested reaction, control experiments in the absence of enzyme resulted in no conversion.

Selbst das Staubkorn sollte uns nicht ärgern,

hat es doch im Universum seinen ganz eigenen Sinn.

Gustav Rädler

CHAPTER 4

EXPANDING THE BIOCATALYTIC TOOLBOX OF FLAVOPROTEIN MONOOXYGENASES FROM *RHODOCOCCUS JOSTII*

This chapter is based on: ***Journal of Molecular Catalysis B: Enzymatic (2013), 88, 20-25*** *by A. Riebel, G. de Gonzalo and M. W. Fraaije.*

Abstract: With the aim to enlarge the set of available flavoprotein monooxygenases, we have cloned eight unexplored genes from *Rhodococcus jostii* RHA1 that were predicted to encode class B flavoprotein monooxygenases. Each monooxygenase can be expressed as soluble protein and has been tested for conversion of sulfides and ketones. Not only enantioselective sulfoxidations, but also enantioselective Baeyer-Villiger oxidations could be performed with this set of monooxygenases. Interestingly, in contrast to known class B flavoprotein monooxygenases, all studied biocatalysts showed no nicotinamide coenzyme preference. This feature coincides with the fact that the respective sequences appear to form a discrete group of sequence related proteins, distinct from the known class B flavoprotein monooxygenases subclasses: the so-called flavin-containing monooxygenases (FMOs), N-hydroxylating monooxygenases (NMOs) and Type I Baeyer-Villiger monooxygenases (BVMOs). Taken together, these data reveal the existence of a new subclass of class B flavoprotein monooxygenases, which we coined as Type II FMOs that can perform Baeyer-Villiger oxidations and accepts both NADPH and NADH as coenzyme. The uncovered biocatalytic properties of the studied Type II FMOs makes this newly recognized subclass of mono-oxygenases of potential interest for biocatalytic applications.

4.1. Introduction

Flavoprotein monooxygenases are attracting attention as selective and oxidative biocatalysts that can be used for the production of high-value chemical building blocks or pharmaceuticals (van Berkel *et al.* 2006). These biocatalysts efficiently catalyze chemo-, regio-, and/or enantioselective oxygenations using dioxygen as mild oxidant while using NAD(P)H as reductant. Flavoprotein monooxygenases can be divided into 6 distinct classes, with each class containing sequence- and structure-related monooxygenases. Two of these classes (class A and B) are especially appealing when considering biocatalysis. These two classes are typified by being single component enzymes that contain a tightly bound flavin cofactor, while the other monooxygenases rely on multiple protomers and often employ a loosely bound flavin cofactor. The class A of flavoprotein monooxygenases seems to have evolved to catalyse aromatic hydroxylations, as most of the characterized monooxygenases of this class represent hydroxylases, typically acting on a very restricted number of substrates (Montersino *et al.* 2011). Class B flavoprotein monooxygenases do not catalyse hydroxylations but perform Baeyer-Villiger oxidations and/or oxygenations of heteroatom-containing compounds. In fact, three class B flavoprotein monooxygenase subclasses have been identified based on specific sequence motifs, which coincide with a preference for specific oxygenation types for each subclass (Fraaije *et al.* 2002):

i) Baeyer-Villiger monooxygenases (BVMOs) contain the sequence motif (FxGxxxHxxxWP/D) and primarily catalyse Baeyer-Villiger oxidations, while they are also able to oxygenate heteroatom-containing compounds (N, S, B or Se containing compounds);

ii) the so-called Flavin-containing monooxygenases (FMOs) contain a slightly different sequence motif (FxGxxxHxxxYK/R), and are specialized in oxidizing heteroatom-containing compounds while they are inefficient in catalyzing Baeyer-Villiger oxidations. The FMOs have mainly been studied as xenobiotic degrading enzymes that help the human body to dispose toxic compounds (Ziegler 1990, Dolphin *et al*. 1997). The human proteome encompasses six FMO isoforms (Hines *et al*. 2002), which are able to activate or degrade many drugs. Only very recently FMOs have been considered for their use as

biocatalysts, due to the identification and production of a microbial FMO which, in contrast to the human homologs, can be easily expressed as a soluble protein (Rioz-Martínez *et al.* 2011).

iii) N-hydroxylating monooxygenases (NMOs) share sequence homology with the above-mentioned class B flavoprotein monooxygenases but lack a typifying sequence motif. Only a conserved histidine can be identified in the region of the BVMO/FMO sequence motif. So far only a few NMOs from bacteria and fungi have been reported. They typically convert long-chain primary amines by *N*-hydroxylation (Stehr *et al.* 1998).

The overall sequence homology among all class B flavoprotein monooxygenases reflects the fact that they are all single-component FAD-containing monooxygenases composed of two dinucleotide binding domains (Rossman folds to bind both FAD and NADPH), that allow them to combine flavin reduction and monooxygenation in one polypeptide chain. They often prefer the use of NADPH as electron donor, keeping the $NADP^+$ tightly bound throughout the catalytic cycle (van den Heuvel *et al.* 2005, Torres Pazmiño *et al.* 2008). From the three subclasses, mainly the BVMOs have been extensively studied as biocatalysts.

To tap the natural diversity for the discovery of novel oxygenating enzymes *Rhodococcus jostii* RHA1 is a very promising candidate. The proteome of this bacterium is predicted to contain an unusually high variety of oxidative enzymes (McLeod *et al.* 2006, Montersino & van Berkel 2012). Inspired by this observation, we and the Grogan group have recently cloned 22 putative BVMO-encoding genes, and succeeded in producing and exploring the biocatalytic properties of these enzymes (Szolkowy *et al.* 2009, Riebel *et al.* 2012). This research has confirmed that all these enzymes indeed act as BVMOs, which can be used for a large number of oxygenations. When screening the predicted proteome of *R. jostii* RHA1 for monooxygenases, we also identified a relatively large number of other putative class B monooxygenases that seem to be more closely related to NMOs and FMOs. While BVMOs are relatively rare enzymes that are only found in bacteria and fungi with an average distribution of only one or two BVMO-encoding genes per microbial genome (de Gonzalo *et al.* 2010), FMOs and NMOs are even more scarce in microbes.

FMOs are quite abundantly present in higher eukaryotes (*e.g.* the 6 isoforms in the human genome and often more than 10 in plant genomes (Schlaich 2007)), but these FMOs are often difficult to produce due to their membrane association.

In this chapter, we report on the exploration of 8 novel class B flavoprotein monooxygenases obtained from a single microorganism, R. jostii RHA1, that share sequence homology with FMOs and NMOs. By optimizing expression vectors and expression conditions, all enzymes were obtained in soluble and holo form. A set of ketones and aromatic sulfides was tested for all enzymes to explore their biocatalytic potential. Also their coenzyme specificity and enantioselective properties were analysed.

4.2. Results

4.2.1. Identification of putative FMO/NMO-encoding genes

By a PBLAST search of the proteome of *R. jostii* RHA1 (McLeod *et al.* 2006), not only 23 Type I BVMO-encoding genes could be found (Riebel *et al*. 2012), but also another eight genes putatively encoding class B flavoprotein monooxygenases could be identified. All corresponding proteins contain two typical Rossman fold motifs (GxGxxG), clearly distinguishing them from other flavoprotein monooxygenase classes. As for all class B flavoprotein monooxygenases, one Rossman fold motif is close to the N-terminus while the other is in the middle of the sequences. In the previous chapter, we proposed a new conserved sequence motif typical for the Type I BVMO family: [A/G]GxWxxxx[F/Y]P[G/M]xxxD, located between the two Rossman fold motifs. This motif entails the conserved active site aspartate and therefore appears a better fingerprint for Type I BVMO sequences. The motif is absent in all sequences used in this study and confirms that they do not represent classical BVMOs. The sequences do also not contain the previously described BVMO-typifying motif (Fraaije *et al.* 2002). In fact, most of them

(all except for monooxygenase H) appear more closely related to FMOs as the FMO-typifying sequence motif can be recognized with only one or two amino acid substitutions (Table 4-1). This is also reflected by the fact that most of the sequences have been annotated in the sequence database as putative flavin-containing monooxygenases. When performing a multiple sequence alignment, the clustering becomes apparent. Monooxygenase H is most related to known NMOs, while the other proteins form another distinct group of sequence-related proteins (Figure 4-1). It also shows that all studied sequences are quite distant from Type I BVMOs. Based on these observations, we annotate monooxygenase H as an NMO (NMO-H). The other proteins seem to form another isolated cluster of sequence-related class B flavoprotein monooxygenases and therefore we classify them as Type II FMOs (while the known and well-established FMO class can be considered as Type I FMOs). Recently another class B flavoprotein monooxygenase has been described that shows relatively high sequence similarity with these Type II FMOs (FMO-X in Figure 4-1). The Grogan group has discovered this bacterial monooxygenase that can perform sulfoxidations and can act as BVMO (Jensen *et al.* 2012). However, conversion rates described for this monooxygenase seem rather modest (100-1000 fold lower when compared with conversion rates of Type I BVMO).

This sequence analysis hints to a new subclass of class B flavoprotein monooxygenases, Type II FMOs that can be employed as biocatalysts. By cloning, expressing, and testing biocatalytic oxidations, we set out to explore the above-mentioned 8 putative class B flavoprotein monooxygenases.

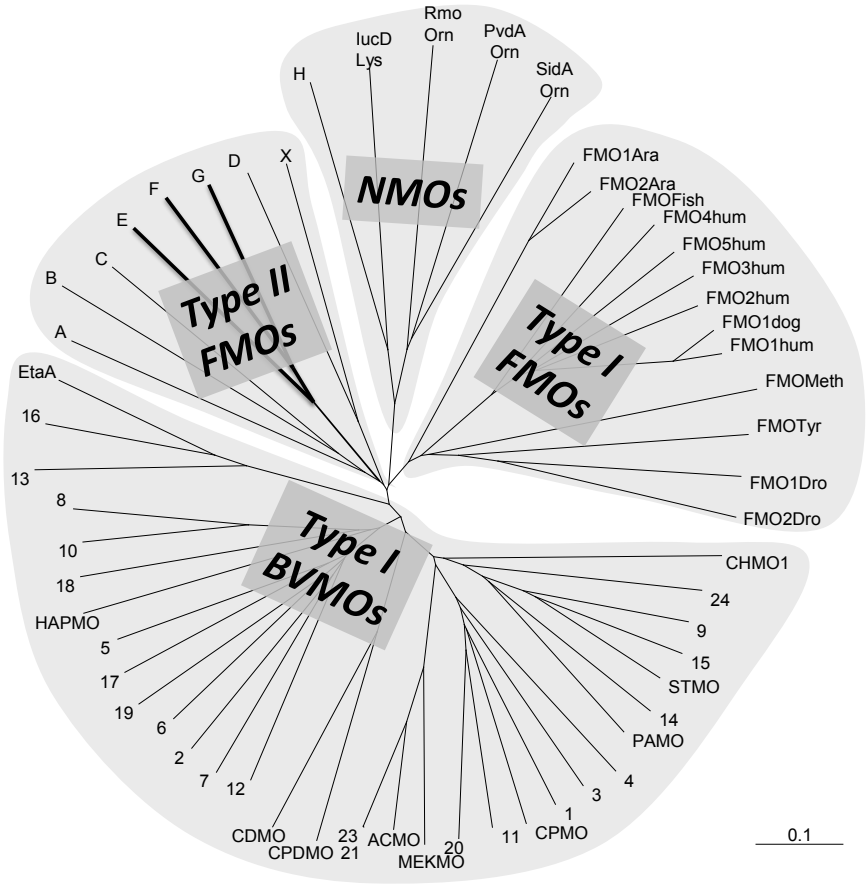


Figure 4-1: An unrooted phylogenetic tree of the studied monooxygenases and known class B flavoprotein monooxygenases. The included sequences are: Type I BVMOs [EtaA, Mycobacterium tuberculosis H37Rv (NP_218371.1); HAPMO, *Pseudomonas fluorescens* ACB (Q93TJ5.1); CDMO, *Rhodococcus ruber* SCI (AAL14233.1); CPDMO, *Pseudomonas sp*. strain HI-70 (BAE93346.1); CHMO1, *Acinetobacter calcoaceticus* NCIMB 9871 (BAA86293.1); PAMO, *Thermobifida fusca* (YP_289549.1); STMO, *Rhodococcus rhodochrous* (BAA24454.1); ACMO, *Gordania sp*. strain TY-5 (BAF43791.1); MEKMO, Pseudomonas veronii MEK700 (ABI15711.1); CPMO, *Comamonas sp*. strain NCIMB 9872 (Q8GAW0)], [BVMOs 1-23 from *Rhodococcus jostii* RHA1 (see [13] for numbering)], Type I FMOs [*Arabidopsis thaliana*: FMO1Ara(Q9LMA1); FMO2Ara (Q9FKE7); *Homo sapiens*: FMO1hum(Q01740); FMO2hum (Q99518); FMO3hum (P31513); FMO4hum (P31512); FMO5hum (P49326); *Drosophila melanogaster*: FMO1Dro (Q9W1E9); FMO2Dro (Q7K3U4); FMO1dog, *Canis familiaris* (Q95LA2); FMOMeth, *Methylophaga sp*. strain SK1 (JC7986); FMOTyr, *Tyria jacobaeae* (D6CHF7); FMOFish, *Oncorhynchus mykiss* (B2LGF9)], NMOs [NRho_Orn, *Rhodococcus jostii* RHA1 (YP_704660.1); PVDA_Orn, *Pseudomonas aeruginosa* (NP_251076); IucD_Lys, *Escherichia coli* (YP_444061.1); SidA_Orn, *Aspergillus fumigatus* (XP_755103.1)], Type II FMOs [FMO_X, *Stenotrophomonas maltophilia* (B2FLR2); the Type II FMOs A-G from *Rhodococcus jostii* RHA1 described in this study].

4.2.2. Expression of the putative class B flavoprotein monooxygenases

The targeted genes were cloned in a pBAD-based expression vector (pBAD*N* by Kamerbeek *et al*. 2004) by ligation free cloning. By varying the temperature and arabinose concentration, optimal expression conditions were determined. By this approach, six proteins could be obtained in soluble form with good expression levels (Table 4-1). Only for two genes, expression did not yield satisfying amounts of soluble protein. We solved this expression problem by cloning the respective genes into the recently developed pCRE2 expression vector (Torres Pazmiño *et al*. 2009). With these constructs we were able to achieve soluble expression of all targeted proteins that were predicted to represent class B flavoprotein monooxygenases.

Interestingly, in two cases, FMO-E and FMO-G, it was found that during expression the growth medium developed a bluish colour, as shown in Table 4-1. We have previously observed this phenomenon when expressing a phenylacetone monooxygenase mutant capable of converting indole into indigo blue (Torres Pazmiño *et al*. 2007). Also the expression of several Type I BVMOs from *R. jostii* RHA1 led to blue pigment formation (Riebel *et al.* 2012). The same phenomenon has been observed when expressing a bacterial FMO in *E. coli* (Choi *et al*. 2003). Apparently, the respective monooxygenases are able to form indigo blue by oxidizing the endogenous indole. The fact that indigo blue formation is only observed upon expression of distinct enzymes, confirms the functional expression of these monooxygenases, in our case of FMO-E and FMO-G.

Table 4-1: All identified class B monooxygenases with their sequence characteristics and expression performance in *E. coli.*

monooxygenase	*accession number*	*Rossman motif* GxGxxG	*FMO motif* $FxGxxxHxxx^{Y}/_{F}{}^{K}/_{R}$	*Rossman motif* GxGxxG	*length (aa)*	*soluble expression* [a]
FMO-A	RHA1_ro00740	GxGxxG	WxGxxxHxxxYR	GxGxxG	375	++[c]
FMO-B	RHA1_ro03334	GxGxxG	AxIxxxHxxxYR	GxGxxG	418	++
FMO-C	RHA1_ro04494	GxGxxG	PxIxxxHxxxYR	GxSxxG	365	++
FMO-D	RHA1_ro05032	GxGxxG	FxGxxxHxxxYS	GxGxxA	369	++[c]
FMO-E [b]	RHA1_ro00824	GxGxxG	FxGxxxHxxxYD	GxGxxA	580	++
FMO-F	RHA1_ro04244	GxGxxG	FxGxxxHxxxHP	GxNxxA	602	+++
FMO-G [b]	RHA1_ro05696	GxGxxG	FxGxxxHxxxFV	GxCxxG	595	+++
NMO-H	RHA1_ro08654	GxGxxN	H	GxGxxG	447	+

[a] **+ indicates a clear visible protein band using SDS-PAGE, while +++ indicates a great overexpression was observed;** [b] **indicates that the medium changed its color to blue during expression of this gene;** [c] **indicates that the respective monooxygenase has been expressed using the pCRE2 expression vector while all others have been expressed using the pBAD*N*-vector (Kamerbeek *et al.* 2004).**

4.2.3. Substrate profile of the overexpressed monooxygenases

Even though using CCE for the reactivity tests, we used the same concentration of recombinant enzyme in all assays to allow a direct comparison of the results. For this, the amount of active enzyme in the CCEs was first quantified with the NADPH-mediated flavin reduction method established and described before (Riebel *et al*. 2012). This revealed that all extracts contained significant amounts of NADPH-dependent flavoproteins, which is in agreement with the SDS-PAGE analysis of CCEs, and confirms that all enzymes contain a flavin cofactor upon expression and accept NADPH as coenzyme.

To determine the biocatalytic potential of the studied monooxygenases, each monooxygenase was tested for activity on several ketones and aromatic sulfides: 2-indanone, cyclopentadecanone, phenylacetone, bicyclo[3.2.0]hept-2-en-6-one, thioanisole, 4-methylthioanisole, ethyl benzyl sulfide, and benzyl phenyl sulfide.

None of the monooxygenases was found to convert 2-indanone or cyclopenta-decanone. In contrast, intriguingly, FMO-E, FMO-F, and FMO-G were found to be able to convert phenylacetone and bicyclo[3.2.0]hept-2-en-6-one (Figure 4-2).

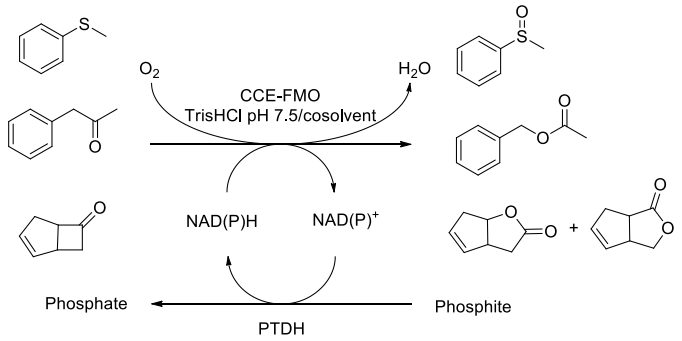


Figure 4-2: Oxidations of thioanisole, phenylacetone and bicycloheptenone catalyzed by the novel flavoprotein monooxygenases using NADH or NADPH as coenzymes.

Product analysis by GC revealed that phenylacetone was converted into the corresponding ester, benzylacetate (Table 4-2). Enantiodivergent oxidation of bicyclo[3.2.0]hept-2-en-6-one was analyzed by chiral GC. This revealed that the monooxygenases have a preference for the formation of the normal lactone, while still a significant amount of the abnormal lactone is formed (Table 4-2). While the biocatalysts are not very enantioselective in formation of the normal lactone, the FMO-F and FMO-G yield >65% ee for the abnormal (1*S*,5*R*) enantiomer. Taken together, these data show that FMO-E, FMO-F and FMO-G act as *bona vide* BVMOs.

Table 4-2: Baeyer-Villiger oxidation of phenylacetone and (±)-bicyclo[3.2.0]hept-2-en-6-one.

	phenylacetone	***(±)-bicyclo[3.2.0]hept-2-en-6-one***			
	conv. (%)	**conv. (%)**	**ratio normal:abnormal**	**ee normal (%)**	**ee abnormal (%)**
FMO-E	**20**	**100**	**79:21**	**9 (1*S*,5*R*)**	**7 (1*R*,5*S*)**
FMO-F	**19**	**88**	**78:22**	**< 1 (1*S*,5*R*)**	**70 (1*S*,5*R*)**
FMO-G	**5**	**100**	**78:22**	**7 (1*R*,5*S*)**	**66 (1*S*,5*R*)**

Next, four prochiral aromatic sulfides with varying substituents (thioanisole, 4-methyl-thioanisole, ethyl benzyl sulfide, and benzyl phenyl sulfide) were tested as substrates. The relatively small substrate thioanisole was the only sulfide that was converted by most of the monooxygenases in detectable amounts, as indicated in Table 4-3. Only NMO-H did

not convert significant amounts of thioanisole or any other tested sulfide, strengthening the classification of this particular enzyme as an NMO.

Table 4-3: Sulfoxidation of thioanisole.

	thioanisole		
	conv. (%)	ee (%)	configuration
FMO-A	5	11	*R*
FMO-B	4	7	*R*
FMO-C	4	33	*R*
FMO-D	16	25	*R*
FMO-E	56	8	*R*
FMO-F	9	≤ 5	-
FMO-G	21	57	*S*

For the majority of the biocatalysts, mainly the (*R*)-methyl phenyl was preferentially formed, with the exception of FMO-G, which led to the (S)-enantiomer. The enantioselectivity of the thioanisole sulfoxidation varied from low to moderate, achieving the highest optical purity (*ee* = 57%), when employing FMO-G. The conversion of 4-methylthioanisole could only be detected in very low conversions. This is in line with the observation of Grogan that FMO-X only showed a poor conversion rate with this aromatic sulfide (Jensen *et al.* 2012). The enantioselective outcome with this substrate when using the Type II FMOs from *R. jostii* RHA1 was similar to that achieved with thioanisole. The more bulky substrate benzyl ethyl sulfide was only converted by FMO-E, FMO-F and FMO-G, but with poor conversion. Only the corresponding methyl phenyl sulfone was detected as product, indicating that these biocatalysts are able to oxygenate the initially formed sulfoxide to the overoxidized product. The bulky benzyl phenyl sulfide was not converted by any of the tested monooxygenases.

FMO-X was found to be indifferent towards NADPH or NADH (Jensen *et al*. 2012). This is in sharp contrast with the known Type I BVMOs, which only accept NADPH as coenzyme. It also is different from the Type I FMOs that typically prefer NADPH as a coenzyme. Conversion of several of the previously identified substrates was therefore

tested via GC with both nicotinamide cofactors. When we replaced NADPH by NADH, almost identical degrees of conversion and enantioselectivities were found for the conversions for all tested Type II FMOs. This is a highly interesting and relevant finding as it suggests that the Type II FMOs have as general characteristic a relaxed coenzyme specificity. This is attractive for biocatalytic applications, as NADH is a cheaper source than its phosphorylated analogue (when considering use as isolated enzyme) while NADH is also present at higher level in cells (when considering use of whole cells). This shows that the Type II FMOs described in this chapter, together with the previously described FMO-X (Jensen *et al.* 2012), represent a newly discovered subclass of class B flavoprotein monooxygenases that can accept both NADPH and NADH.

4.3. Discussion

Eight class B flavoprotein monooxygenases from *R. jostii* RHA1 have been expressed in soluble form and studied for their biocatalytic potential. From the sequence alignment study, one of the targeted monooxygenases (NMO-H) was found to belong to the group of NMOs. The prototype enzymes for the subclass of *N*-hydroxylating flavoprotein monooxygenases (NMOs) are L-ornithine N5-hydroxylase and L-lysine N6-hydroxylase (Olucha & Lamb 2011). An L-ornithine hydroxylase had already been identified in the *R. jostii* RHA1 proteome (Rmo_Orn in Figure 4-1, Bosello *et al.* 2012). When comparing NMO-H with known NMO sequences it shows highest sequence similarity with the lysine hydroxylase IucD from *E. coli* (Figure 4-1). This suggests that the NMO-H described in this study probably represents a lysine hydroxylase present in *R. jostii* RHA1. The restricted substrate scope of lysine hydroxylases also is in line with the observation that NMO-H is not converting any aromatic compound tested in this study (Franceschini *et al.* 2012). This confirms that, in contrast to BVMOs and FMOs, NMOs are rather restricted in substrate acceptance and appear to be of little value as biocatalysts.

The FMOs described in this chapter do not seem to belong to the typical FMO subclass, as (i) they do not contain a fully conserved FMO-typifying motif (Table 4-1), and (ii) they form another cluster of sequence-related proteins distinct from the three other

subclasses, FMOs, NMOs, and BVMOs (Figure 4-1). Also their catalytic properties deviate from the other subclasses by accepting both nicotinamide coenzymes, NADH and NADPH, without significant difference. Furthermore, several of these FMOs were shown to perform Baeyer-Villiger oxidations, a reaction for which typical FMOs display hardly any activity. We coined this newly recognized subclass of class B flavoprotein monooxygenases Type II FMOs.

This study reveals that of the tested Type II FMOs, FMO-E, FMO-F, and FMO-G may represent valuable new oxidative biocatalysts. These three enzymes were found to be active in both sulfoxidations and Baeyer-Villiger reactions. Multiple sequence alignment revealed that they form a separate cluster of sequences (Figure 4-1). Except for showing a relative high sequence homology with each other, these three monooxygenases also share an *N*-terminal extension of about 160 residues (see multiple sequence alignement in Appendix). This may suggest that this newly identified subclass of monooxygenases has evolved as separate subclass of class B flavoprotein monooxygenases, distinct from the Type I BVMOs, towards a new subclass of potent Baeyer-Villiger monooxygenases. Therefore, these three FMOs or close homologs thereof may represent interesting alternative biocatalysts (for instance, due to their NADH acceptance) when compared with the known Type I BVMOs. By merely using the N-terminal sequence of one of the three Type II FMOs, it is possible to identify hundreds of homologous putative Type II FMOs. It will be exciting to see whether these homologs indeed display similar biocatalytic properties. Future research will reveal whether monooxygenases from this newly identified subclass of class B flavoprotein monooxygenases (Type II FMOs) are indeed valuable as biocatalysts; *e.g.* their operational stability and substrate scope has still to be explored.

In the present chapter, a set of eight new flavoprotein monooxygenases from R. jostii RHA1 has been expressed and obtained in soluble form by employing two expression vectors and optimizing expression conditions. Sequence analysis

established that one of these enzymes represents an NMO, while the other cluster as a separate class B flavoprotein monooxygenase subclass: Type II FMOs. The monooxygenases have been tested for their ability to perform sulfoxidations and Baeyer-Villiger oxidations. Interestingly, three of the Type II FMOs are able to catalyze both reaction types and accept both NADPH or NADH as coenzyme, presenting a different behavior with respect to classical FMOs and BVMOs. Type II FMOs might develop as an interesting alternative to BVMOs, as they are able to catalyze the similar reaction with a relaxed nicotinamide cofactor acceptance.

4.4. Experimental Section

General materials and methods

Oligonucleotide primers were purchased from Sigma, dNTPs and In-Fusion™ 2.0 CF Dry-Down PCR Cloning Kit from Clontech, Phusion polymerase from Finnzymes. All other chemicals were obtained from Acros Organics, ABCR, Sigma-Aldrich, TCI Europe, and Roche Diagnostics GmbH. The nicotinamide coenzymes were purchased from Codexis.

The NCBI server was used for DNA sequence retrieval and BLAST searches. The EBI server was used for multiple sequence alignment by CLUSTALW (Thompson *et al.* 1994). Treeview software was used for visualization of the sequence relationships.

NCBI server: http://blast.ncbi.nlm.nih.gov/Blast.cgi

EBI server: http://www.ebi.ac.uk/Tools/msa/clustalw2/

Treeview software: http://taxonomy.zoology.gla.ac.uk/rod/treeview.html

Cloning and expression

Escherichia coli TOP10 from Invitrogen was used as a host for DNA manipulations and protein expression. Two expression vectors have been used: (*i*) a modified pBAD vector (pBAD*N*) in which the NdeI site was replaced by the original NcoI site (Kamerbeek *et al.* 2004), and (*ii*) the recently engineered pBAD-based pCRE2 vector which harbors a codon-optimized gene encoding a thermostable mutant of phosphite dehydrogenase (PTDH) with an N-terminal His-tag (Torres Pazmiño *et al*. 2009).

The target genes were amplified by PCR using genomic DNA of *R. jostii* RHA1 as template and subsequently cloned into pBAD*N* using the In-Fusion PCR Cloning kit from Clontech, following the recommendations of the manufacturer. Expression of all generated expression constructs was tested using 24-multiwell microtiter plates in the sandwich cover system from Enzyscreen®. Cell cultures of 2,5 ml in LB medium supplemented with 50 µg mL^{-1} ampicillin were grown at four different temperatures (17°C for 48 h; 24°C for 32 h; 30°C and 37°C for 16 h) with four different arabinose concentrations (0,002%, 0,02%, 0,2%, and none) each, with shaking at 200 rpm. Cell extracts (CEs) were obtained using DNase/lysozyme in combination with freezing in liquid nitrogen and thawing at 30°C. CEs as well as cell cleared extracts (CCEs) were analyzed by SDS-PAGE for (soluble) expression of the monooxygenases. For the genes that did not yield expressed soluble protein when cloned into pBAD*N*, the in-house developed pCRE2 expression vector was used.

General procedure for the bioconversions employing the novel FMOs

Conversions were performed essentially as described before (Riebel *et al.* 2012). For GC analysis, 500 µl incubations of 50 mM Tris/HCl pH 7,5, 10% glycerol, 1 mM DTT, 1 mM EDTA, 10 µM FAD, 5 mM substrate, 5% cosolvent (1,4-dioxane), 100 µM NADPH, 3,1 µM PTDH, 10 mM phosphite and 5 µM of the corresponding monooxygenase in CCE form (total volume of 0,5 mL), were shaken in glass vials at 24°C for 24 hours. For determining the exact concentration of each enzyme in the respective extract, a recently developed method was used which relies on the decrease in absorbance at 450 nm upon NADPH-mediated reduction of the flavin cofactor (Riebel *et al.* 2012). Each conversion was stopped by extracting with ethyl acetate (2 x 0,5 mL containing 0,1% mesitylene as standard), dried with magnesium sulfate and analyzed directly by GC to determine the degree of conversion and the enantioselectivity. The details concerning the (chiral) GC analysis can be found in the Appendix A-1. For every tested reaction, control experiments in the absence of enzyme resulted in no conversion.

Wunder stehen nicht im Gegensatz zur Natur,
sondern nur im Gegensatz zu dem,
was wir über die Natur wissen.
St. Augustin

CHAPTER 5

TYPE II FLAVIN-CONTAINING MONOOXYGENASES: A NEW CLASS OF BIOCATALYSTS THAT HARBORS BAEYER-VILLIGER MONOOXYGENASES WITH RELAXED COENZYME SPECIFICITY

This chapter is based on ***chemcatchem (2013), accepted*** *by Anette Riebel, Michael J. Fink, Marko D. Mihovilovic, Marco W. Fraaije.*

Abstract: Within a newly identified subclass of flavin-containing monooxygenases (FMOs) we have identified three monooxygenases from *Rhodococcus jostii* RHA1 (FMO-E, FMO-F, and FMO-G) that are effective in catalyzing Baeyer-Villiger oxidations. All three Type II FMOs display a relaxed coenzyme specificity by accepting both NADPH and NADH as coenzyme, a novel and attractive feature among biocatalysts capable to conduct Baeyer-Villiger oxidations. We have purified FMO-E and have determined that the K_M values for both coenzymes are in the micromolar range while the activity is highest for NADH. By using the stopped-flow technique, formation of a relatively stable peroxyflavin-enzyme intermediate was observed which indicates that Type II FMOs employ a catalytic mechanism which is similar to other class B flavoprotein monooxygenases. A large set of cyclobutanones and cyclohexanones was used to probe the regio- and enantioselectivity of all three recombinant monooxygenases. The biocatalysts readily accept a large number of cyclobutanones, enabling conversion of previously poorly accepted substrates by other mono-oxygenases (esp. norcamphor), and exhibit excellent and unique regio- and enantioselectivities. Sequence analysis revealed that Type II FMOs that act as Baeyer-Villiger monooxygenases contain a unique N-terminal domain. Sequence conservation in this protein domain can be used to identify new NADH-dependent Baeyer-Villiger mono-oxygenases, facilitating future biocatalyst discoveries.

5.1. Introduction

By being able to perform efficient chemo-, regio-, and/or enantioselective oxygenation reactions by only using dioxygen and reducing equivalents like NADH and NADPH (Torres Pazmiño *et al.* 2010), flavoprotein monooxygenases represent potent catalytic tools for the production of high-value chemical building blocks or pharmaceuticals (van Berkel *et al.* 2006). Flavoprotein monooxygenases can be classified into six structurally and mechanistically distinct classes. When taking into account the differences in structural and mechanistic properties of the various flavoprotein monooxygenase classes, class A and B monooxygenases are especially interesting for biocatalysis as they are single component enzymes with a typically tightly bound FAD cofactor (Figure 5-1).

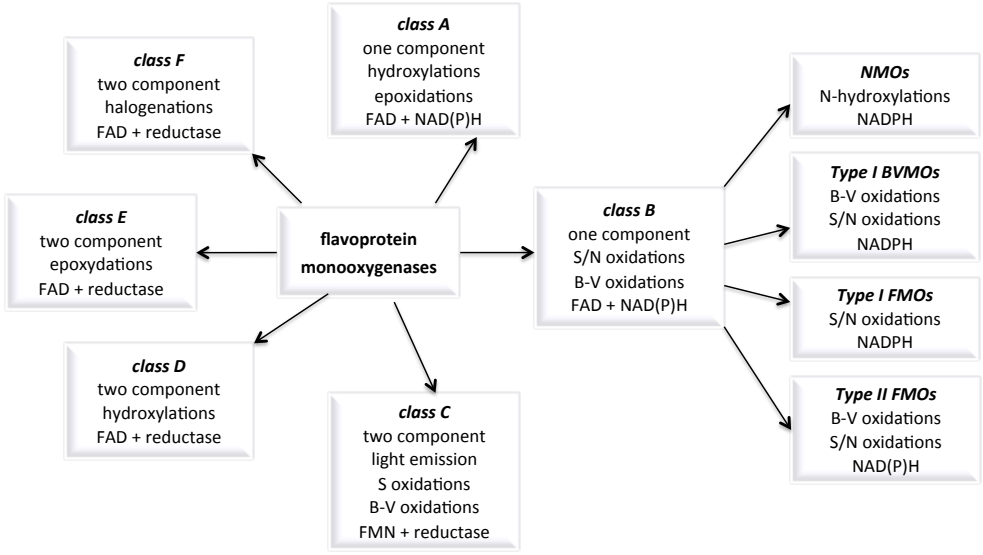


Figure 5-1: Overview of the different class B flavoprotein monooxygenase classes; S/N-oxidations: sulfur and nitrogen oxidations; B-V oxidations: Baeyer-Villiger oxidations.

Class A flavoprotein monooxygenases have been shown to be specialized in hydroxylation and epoxidation reactions (van Berkel *et al*. 2006). Class B flavoprotein monooxygenases typically cover another range of oxygenation reactions, such as Baeyer-Villiger oxidations and oxygenations of heteroatom-containing compounds (van Berkel *et al*. 2006). Careful

analysis of the sequence and enzymatic properties of class B flavoprotein monooxygenases has revealed several distinct subclasses (Figure 5-1).

An extensively studied subclass of class B flavoprotein monooxygenases is formed by the Type I Baeyer-Villiger monooxygenases (BVMOs): they display unique catalytic properties by catalysing NADPH-dependent Baeyer-Villiger oxidations, while they are also capable of oxygenating heteroatom-containing compounds (de Gonzalo *et al*. 2010). Three other subclasses of class B flavoprotein monooxygenases are known to date. The Type I flavin-containing monooxygenases (FMOs) have mainly been studied for their role in human metabolism (Ziegler 1990). The human proteome contains six Type I FMO isoforms, all of which have been reported as important in the oxidative degradation of N- and S-containing compounds. Only recently, some microbial representatives have been discovered and were shown to represent interesting biocatalysts for oxidizing secondary amines, including the conversion of indole into indigo blue (Rioz-Martínez *et al*. 2011). Another class B subclass entails microbial monooxygenases that hydroxylate primary amines, *e.g.* lysine N6-hydroxylase and ornithine N5-hydroxylase. Referring to their reactivity, these have been named N-hydroxylating monooxygenases (NMOs). All together, these biocatalyst subclasses share a pronounced cofactor preference towards NADPH.

In the last chapter, we described a new group of sequence-related monooxygenases that belong to the class B flavoprotein monooxygenases: Type II FMOs (Figure 5-1). We reported on seven Type II FMOs derived from the genome of Rhodococcus jostii RHA1 (Riebel et al. 2013). Their unique sequence profiles as well as their relaxed coenzyme specificity legitimated a classification separate from the other class B flavoprotein monooxygenases. Substrate profiling of the seven rhodococcal mono-oxygenases revealed that three of them are able to perform Baeyer-Villiger oxidations. In this chapter, we report on the characterization of these three Type II FMOs that can act as bona fide BVMOs.

5.2. Results

5.2.1. Purification of FMO-E

We have recently identified seven Type II FMO-encoding genes in the proteome of *R. jostii* RHA1. The respective enzymes, together with the monooxygenase of *Stenophomamonas maltophilia* (Jensen *et al*. 2012), are members of a new subclass of class B flavoprotein monooxygenase: a class of sequence-related FAD-containing monooxygenases with a relaxed coenzyme preference towards NADPH and NADH (Riebel *et al.* 2013). Within these seven newly found Type II FMOs, however, we found significant differences concerning substrate acceptance. The monooxygenases FMO-E, FMO-F, and FMO-G efficiently catalyzed Baeyer-Villiger oxidation of two ketones (phenylacetone (substrate **3**) and bicyclo[3.2.0]hept-2-en-6-one (substrate **1**)); a catalytic feature that was only known for Type I BVMOs. However, different from Type I BVMOs, all Type II FMOs were shown to be able to use NADPH or NADH without any preference. This aspect seems particularly interesting for combining such enyzmes with other redox biocatalysts in multi-step biotransformations cascades. These features encouraged us to perform a more in-depth study of the newly identified Type II FMOs exhibiting BVMO activity.

To allow efficient one-step purification, we produced the monooxygenases FMO-E, FMO-F, and FMO-G with an N-terminal Strep-tag (in pBAD*NS*). The strep-tag proved to have no influence on the expression level. However, purification trials revealed that only FMO-E could be isolated in good yields. The other monooxygenases were found to suffer from FAD loss during the purification procedure. As all three monooxygenases are sequence related, we decided to use FMO-E as prototype for a biochemical characterization. From a 1 L culture, 25 mg pure FMO-E could be obtained (see Appendix). The purified enzyme displays a typical flavin absorbance spectrum, with characteristic absorbance maxima at 451 nm and 381 nm (Figure 5-2). The extinction coefficient was determined as 11,9 $mM^{-1}cm^{-1}$ (ε_{451nm}). Furthermore, from the measured A280/A451 ratio of 10 it can be concluded that the enzyme is purified as holo enzyme.

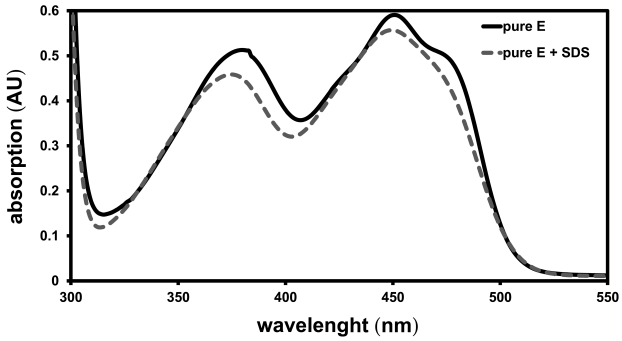


Figure 5-2: Absorption spectra of purified native FMO-E (line) and after SDS treatment (dotted line).

5.2.2. Kinetic characterization of FMO-E

Using purified FMO-E, the steady-state kinetic parameters were determined for bicyclo[3.2.0]hept-2-en-6-one (**1**), NADH and NADPH. For both coenzymes and the ketone substrate typical Michaelis-Menten behavior was observed, yielding the catalytic (k_{cat}) and Michaelis constants (K_M) for all three substrates (Table 5-1).

Table 5-1: Steady-state kinetic parameters of FMO-E for NADPH, NADH and bicycloheptenone.

substrate	K_M *(mM)*	k_{cat} *(s^{-1})*
bicyclo[3.2.0]hept-2-en-6-one (1)	**19,8**	**2,7**
NADPH	**< 0.0010**	**2,7**
NADH	**0,0054**	**4,3**

Even though FMO-E accepts both nicotinamide coenzymes, the enzyme appears to have a higher affinity (K_M) for NADPH. Yet, the K_M for NADH is also in the low µM range and the k_{cat} is higher when compared with NADPH (4,3 s^{-1} vs. 2,7 s^{-1}). This indicates that FMO-E is an efficient enzyme with NADH as coenzyme, which is in contrast to most other class B flavoprotein monooxygenases. Also all other Type II FMOs have shown coenzyme indifference (Riebel *et al.* 2013). The determined k_{cat} values also confirm FMO-E as a potent biocatalyst as the rate of catalysis is in the same range as it has been found for other class B flavoprotein monooxygenases. For example, the k_{cat} values for the typical Type I BVMOs phenylacetone monooxygenase (Fraaije *et al.* 2005) and cyclohexanone monooxygenase (Sheng *et al.* 2001) are 1-20 s^{-1}. It is particularly noteworthy that the

extremely low k_{cat} values determined for the Type II FMO from *Stenophomamonas maltophilia* (Jensen *et al.* 2012) differ by 2 orders of magnitude relative to FMO-E. This makes FMO-E a more likely candidate for biocatalytic processes.

To establish whether FMO-E utilizes a similar catalytic mechanism when compared with Type I FMOs and Type I BVMOs, stopped-flow experiments were performed. The kinetic mechanism of flavoprotein monooxygenases can be subdivided into two half-reactions: a reductive and an oxidative half-reaction. The reduction of the FAD cofactor by the nicotinamide coenzyme is considered as the reductive half-reaction. The anaerobic reduction of FMO-E by NADPH was monitored over time using a stopped-flow instrument, revealing a relatively high rate of reduction (17,5 s^{-1}) at a saturating NADPH concentration. In the next step, the reduced flavin of class B flavoprotein monooxygenases is able to swiftly react with molecular oxygen, typically yielding a relatively stable peroxyflavin intermediate. The ability to stabilize such a reactive enzyme intermediate in the absence of a substrate has been the typical hallmark of class B flavoprotein monooxygenases. For monitoring the above-described oxidative half reaction, the NADPH-reduced FMO-E was rapidly mixed with aerated buffer. Analysis of the obtained spectral scans revealed, that the data could be fitted best with an irreversible three-step model (A → B → C → D) (Figure 5-3).

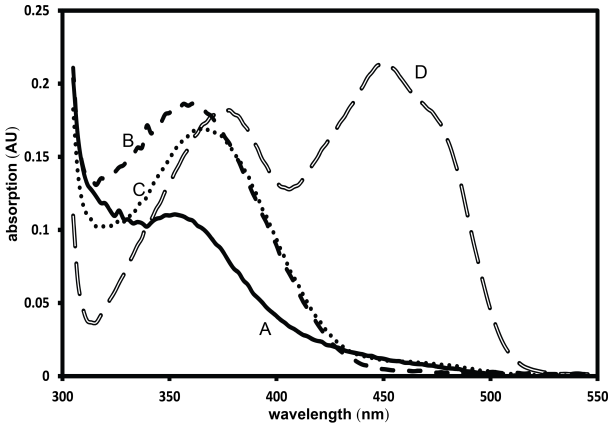


Figure 5-3: Deconvoluted flavin absorbance spectra upon mixing FMO-E with air-saturated buffer. The spectra were reconstructed from a model A to B to C to D and with rates of respectively 7,2, 1,2 and 0,5 s^{-1}. Spectra: A: reduced state, B: first peroxyflavin intermediate, C: second peroxyflavin intermediate, D: oxidized state.

Directly after mixing with oxygen, species B is rapidly formed with a rate of 7,2 s^{-1} as evidenced by an absorbance peak at around 365 nm. This spectrum perfectly matches the spectral features of a peroxyflavin intermediate and confirms that FMO-E operates via a catalytic mechanism that is similar to other class B flavoprotein monooxygenases. The peroxyflavin intermediate (spectrum B) undergoes a slow and modest spectral change (1,2 s^{-1}) to form a less intense absorbance peak (spectrum C). This transition may represent a change in the protonation state of the peroxyflavin or a conformational change, and has also been observed for Type I BVMOs (Mirza *et al.* 2009, Orru *et al.* 2011). In a subsequent slow process (0,5 s^{-1}), the oxidized flavin is formed (spectrum D), accompanied by formation of hydrogen peroxide. These kinetic and spectral properties closely resemble the kinetic behavior of other class B flavoprotein monooxygenases.

5.2.3. Structural analysis of FMO-E

To obtain a better view on how FMO-E is able to catalyze Baeyer-Villiger oxidations, we compared a structural model of FMO-E with the crystal structures of a Type I BVMO (PDB 1W4X) (Malito *et al.* 2004) and a Type I FMO (PDB 2VQ7) (Alfieri *et al*. 2008). Comparison of the active sites clearly showed that all three enzymes employ different strategies to catalyse oxygenations (Figure 5-4).

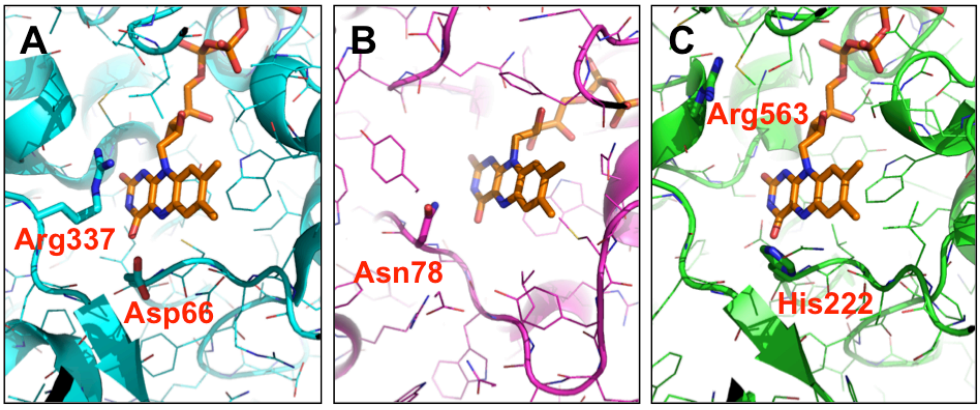


Figure 5-4: Comparison of the active side architectures of (A) a Type I BVMO (PDB:1W4X), (B) a Type I FMO (2VQ7) and (C) FMO-E.

In Type I BVMO a strictly conserved arginine is crucial for catalysis, in Type I FMO an asparagine residue serves such a role, while in the model of FMO-E no such active residue can be identified in the homologous structural positions. Yet, a histidine and an arginine residue were found to be positioned close to the isoalloxazine moiety of the FAD cofactor, possibly assisting in catalysis. Both residues are conserved among Type II FMO sequences. To verify whether His222 and Arg563 serve a role in catalysis, the corresponding alanine mutants were prepared. The first evidence of changed catalytic capacities of the mutants could directly be seen when growing cultures under optimal expression conditions. For the wild type FMO-E it was observed that the growth medium changed significantly in color. This is due to blue pigment formation which is formed upon conversion of indigenous indole into indigo blue by FMO-E (Riebel *et al.* 2013). Expression of the mutants did not result in any color change of the medium. Furthermore, purification of the mutant proteins only resulted in inactive apo protein. Addition of FAD to the apo protein samples did not restore any activity. Tests with cleared cell extracts containing the different mutant enzymes showed that the activity was affected dramatically. The conversion of thioanisole (substrate **2**) dropped from 56% (Riebel *et al*. 2013) to only 6%, while phenylacetone (**3**) and bicyclo[3.2.0]hept-2-en-6-one (**1**) were not converted at all under standard assay conditions.

These results show that the targeted residues are important for proper FAD binding while they do not provide conclusive evidence on whether they are important for catalysis.

5.2.4. Biocatalytic exploration

As Type II FMOs are fundamentally different when compared to Type I BVMOs, we set out to explore the substrate acceptance range and stereoselective properties of all three studied FMOs by using whole cells. For this, 14 ketones (six prochiral cyclobutanones **4-9**, four prochiral cyclohexanones **10-13**, three racemic fused cyclobutanones **1, 14, 15** and rac. norcamphor **16**) were tested as substrates (see Table 5-3 and Table 5-2 for their structural formulas).

Desymmetrization of prochiral cyclobutanones and cyclohexanones revealed a clear preference (Table 5-3). All three monooxygenases efficiently convert most cyclobutanones while the cyclohexanones are poorly or not accepted as substrates. The corresponding lactones are formed with modest to good enantioselectivity (e.g. ranges from 5 – 79 %, median 45 %). Interestingly, Type II FMOs display complementary enantio-preference in all cases, hence, enabling the synthesis of both optical antipodes in enantioenriched form.

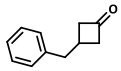
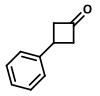
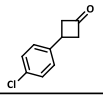
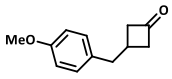
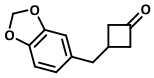
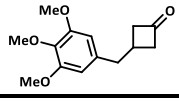
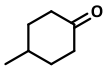
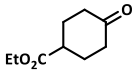
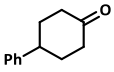
As further exploration of the biocatalytic potential of these newly found monooxygenases, we decided to test conversion of racemic fused cycloketones. Gratifyingly, the monooxygenases were found to be able to convert all tested fused cyclobutanones and norcamphor (Table 5-2). Regiodivergent oxidations of fused cycloketones resulted in preferential formation of the normal lactones (in this case, the rearrangement is governed by higher nucleophilicity, *i.e.* higher degree of substitution of the migrating carbon).

Table 5-2: Regiodivergent oxidation of fused cyclobutanones and norcamphor by FMO-E, FMO-F, and FMO-G using whole cells.

#	Substrate structure	FMO-E			FMO-F			FMO-G		
		% Rel. Conv.	% ee N	E_N	% Rel. Conv.	% ee N	E_N	% Rel. Conv.	% ee N	E_N
		Ratio N:*ABN*	*% ee ABN*	E_{ABN}	Ratio N:*ABN*	*% ee ABN*	E_{ABN}	Ratio N:*ABN*	*% ee ABN*	E_{ABN}
1		97	46 (3a*R*,6a*S*)	n.d.	92	19 (3a*S*,6a*R*)	n.d.	81	21 (3a*S*,6a*R*)	n.d.
		14 : 86	*3 (3aR,6aS)*	*n.d.*	12 : 88	*6 (3aR,6aS)*	*n.d.*	20 : 80	*rac.*	*n.d.*
14		81	39 (3a*S*,7a*S*)	n.d.	77	44 (3a*S*,7a*S*)	137	91	18 (3a*S*,7a*S*)	n.d.
		95 : 5	*98 (3aR,7aS)*	*n.d.*	92 : 8	*71 (3aR,7aS)*	*59*	92 : 8	*75 (3aR,7aS)*	*n.d.*
15		>99	rac.	n.d.	77	91[a]	99	99	rac.	n.d.
		>99:*1*	*n.a.*	*n.a.*	>99:*1*	*n.d.*	*n.d.*	>99:*1*	*n.a.*	*n.a.*
16		14	94 (1*S*,5*R*)	36	22	>99 (1*S*,5*R*)	>200	47	99 (1*S*,5*R*)	>200
		>99:*1*	*n.a.*	*n.a.*	>99:*1*	*n.a.*	*n.a.*	>99:*1*	*n.d.*	*n.a.*

[a] **Product regioisomers are not separable by preparative low-pressure chromatography; for assignment of regio- and stereoisomers see Mihovilovic *et al*. 2008. Relative conversion (Rel. Conv.) and enantiomeric excess (ee) of lactone products as determined by chiral phase GC-FID after 24 h biotransformation time. Selectivity values E in kinetic resolution-type reactions were calculated according to Sih's equation E = [ln(1 – eeS) – ln(1 + eeS/eeP)] / [ln(1 + eeS) – ln(1 + eeS/eeP)] from ee of substrate (eeS) and ee of product (eeP) values. Values for normal lactones (N) in plain font, values for abnormal lactones (ABN) in italics. Absolute configuration or sign of optical rotation of lactone products was assigned on the basis of published reference biotransformations (Mihovilovic *et al.* 2008). n.a.: not applicable, n.d.: not determined, rac.: racemic.**

Table 5-3: Desymmetrization of cyclobutanones and cyclohexanones by FMO-E, FMO-F, and FMO-G using whole cells.

substrate	structure	conversion (%) *ee* lactone (%) E	F	G
4		100 16 (S)	90 40 (R)	95 50 (R)
5		100 69 (R)	96 78 (S)	93 67 (S)
6		100 79 (R)	76 64 (S)	67 47 (R)
7		98 5 (S)	75 18 (S)	99 42 (R)
8		73 30 (R)	83 50 (S)	40 40 (S)
9		10 13 (R)	0,2 n.a.	0 n.a.
10		1 n.a.	0,3 n.a.	0,3 n.a.
11		12 99 (+)[a]	0,2 n.a.	0 n.a.
12		0 n.a.	0,1 n.a.	0,1 n.a.
13		0 n.a.	0,1 n.a.	0 n.a.

[a] **absolute configuration of product not known; sign of optical rotation is given instead. Relative conversion (rel. conv.) and enantiomeric excess (ee) of lactone products as determined by chiral phase GC-FID after 24 h biotransformation time. Absolute configuration or sign of optical rotation of lactone products was assigned on the basis of published reference biotransformations (Rudroff *et al.* 207). n.a.: not applicable.**

This kind of regioselectivity is known from Type I BVMOs of the CPMO cluster (Mihovilovic et al. 2008). Yet, interestingly, these FMOs are the only known biocatalysts able to perform the kinetic resolution of norcamphor (**16**) to the normal lactone with both excellent stereoselectivity and yield. The best reported biocatalyst for this transformation, CDMO from *Rhodococcus ruber* SC1 (Fink *et al.* 2012), catalyzes the formation of both regioisomers of norcamphor with excellent enantioselectivity (>98%), but in almost equimolar ratio (57:43 normal/abnormal) and at a rather low rate. Specifically, FMO-F and FMO-G display excellent enantioselectivities (E >200) and regioselectivities (>99:1 for normal:abnormal lactone) for the oxygenation of this natural compound.

Clearly, the studied Type II FMOs display biocatalytic features that set them apart from Type I BVMOs. As a consequence, they represent a new and promising group of oxidative biocatalysts.

5.3. Discussion

Within the newly described class of Type II FMOs we have identified three FMOs that can efficiently catalyze Baeyer-Villiger oxidations. The kinetic analysis of one of these monooxygenases, FMO-E, has revealed that it is catalytically competent by displaying high affinities towards both nicotinamide coenzymes and high k_{cat} values. The K_M values for NADPH and NADH are in the low µM range while the highest rate of catalysis is observed with NADH (4,3 vs. 2,7 s^{-1}). These rates are significantly higher when compared with a recently reported Type II FMO from *S. maltophilia* that can also perform Baeyer-Villiger oxidations (0,029 s^{-1}). Interestingly, for both bacterial monooxygenases the rate of catalysis is highest when using NADH. This may hint to a role of the coenzyme in the rate-limiting step of the kinetic mechanism of this monooxygenase. For the Type I BVMO cyclohexanone monooxygenase it has indeed been shown that $NADP^+$ release limits the rate of catalysis (Sheng *et al.* 2001).

The pre-steady state kinetic analysis of FMO-E has revealed similar features concerning the catalytic mechanism as with other class B monooxygenases: (*i*) oxidized FMO-E is rapidly reduced by a reduced coenzyme without the need of substrate binding,

(*ii*) the reduced FMO-E rapidly reacts with molecular oxygen, and (*iii*) FMO-E is able to form a stable peroxyflavin intermediate. For all other studied class B monooxygenases, it has been shown that the oxidized nicotinamide coenzyme remains bound to the active site until the peroxyflavin has oxygenated the substrate. Only then it is released, often in a relatively slow kinetic event. Such a scenario fits with the steady-state and pre-steady-state kinetic data of FMO-E in which $NADP^+/NAD^+$ is released in a relatively slow process. More detailed kinetic and structural studies will settle this issue in the future.

As the studied monooxygenases FMO-E, FMO-F, and FMO-G can act as BVMOs and can use NADH as coenzyme, they represent highly sought-after biocatalysts. A major disadvantage of the well-studied Type I BVMOs is that they are highly specific for NADPH, which is a relatively expensive cofactor when compared with NADH. Attempts to engineer a Type I BVMO that accepts NADH as cofactor have failed so far (Kamerbeek *et al.* 2004, Dudek *et al.* 2010). A relaxed nicotinamide cofactor acceptance is also a highly appealing feature for designing enzyme cascade reactions that involve multiple redox biocatalysts. These monooxygenases will be able to tap from both nicotinamide coenzyme pools in a cell. When employing whole-cell processes, a significantly better balanced exploitation of the available reduction equivalents within a metabolically active cell may be achieved, alleviating the metabolic burden of such systems. Additionally, higher compatibility and flexibility may be achieved when designing redox cascades in a cell-fee environment.

This study shows that the tested Type II FMOs can indeed be used as efficient biocatalysts. By testing a set of cyclobutanone derivatives we have unveiled new and excellent stereoselectivities. For example, the conversion of norcamphor (**16**) by FMO-F proceeds with >99% regioselectivity and >99% enantioselectvity. Such selectivity has not been observed before with previously identified BVMOs. This may be explained by the fact that Type II FMOs are only distantly related to Type I BVMOs. As a consequence they will exhibit totally different active site architectures. In fact, Type II FMOs seem to lack the typical active residues that are employed in Type I BVMOs (a conserved arginine) and Type I FMOs (a conserved asparagine) for catalysis. By site-directed mutagenesis we have determined that His222 and Arg563 are crucial for corect functioning of FMO-E.

Unfortunately the protein sequences of FMO-E and the Type II FMO from *S. maltophilia* (Jensen *et al*. 2012) are too dissimilar to locate the mutated residues in the structure of the latter monooxygenase. Elucidation of the structure of FMO-E would shed light on the structural and catalytic properties of this NAD(P)H-dependent monooxygenase. Intriguingly, on protein sequence, the three special Type II FMOs described in this chapter share an extra N-terminal extension of about 160 residues when compared with the other four rhodococcal Type II FMOs and the Type II FMO from *S. maltophilia*. This aberrant sequence feature coincides with the potency to efficiently catalyze Baeyer-Villiger oxidations. The molecular basis for this correlation is unclear. Yet, it may be used to identify new Type II FMOs that can be exploited as NADH-driven biocatalysts for Baeyer-Villiger oxidations. A BLASTP search in the genome sequence database search reveals that there is a large number of Type II FMOs that contain this typical N-terminal extension. Intriguingly, the respective genes are mainly found in fungal genomes.

This study has shown that Type II FMOs represent valuable oxidative biocatalysts. Especially the ability to use NADH as cofactor opens up new possibilities for cost-effective biocatalytic Baeyer-Villiger oxidations. Many more Type II FMOs await biocatalytic exploration and future research will reveal how they operate at molecular level.

5.4. Experimental Section

Reagents & enzymes

Oligonucleotide primers were purchased from Sigma, dNTPs and In-Fusion™ 2.0 CF Dry-Down PCR Cloning Kit from Clontech, Phusion polymerase from Finnzymes. All other chemicals were obtained from Acros Organics, ABCR, Sigma-Aldrich, TCI Europe, and Roche Diagnostics GmbH.

Bacterial strains and Plasmids

Escherichia coli TOP10 from Invitrogen was used as a host for DNA manipulations and protein expression. The used expression vector was a modified pBAD vector (pBAD*N*) in which the *Nde*I site was replaced by the original *Nco*I site (Kamerbeek *et al.* 2004). For this

study we have introduced a *strep*-tag binding site (tggagccacccgcagtttgaaaaa) followed by a TEV protease cleavage site (gagaatttatattttcaaggt) between promotor and start codon of the gene to be cloned, yielding the pBAD*NS* vector.

Sequence analysis, cloning, expression and protein purification

The NCBI server (blast.ncbi.nlm.nih.gov/Blast.cgi) was used for DNA sequence retrieval and BLAST searches. The EBI server (www.ebi.ac.uk/Tools/msa/clustalw2/) was used for multiple sequence alignment by CLUSTALW (Thompson *et al.* 1994). The FMO-E sequence was used as input to prepare a structural model by using the CPH modeling server (www.cbs.dtu.dk/services/CPHmodels/) (Nielsen *et al.* 2010). The model was made based on the phenylacetone MO structure (PDB 1W4X).

The target genes were amplified by PCR using genomic DNA of *R. jostii* RHA1 as template and subsequently cloned into pBAD*NS,* using the In-Fusion PCR Cloning kit from Clontech, following the recommendations of the manufacturer. The mutations were introduced via the QuikChange® Site-Directed Mutagenesis Kit from Stratagene, following the recommendations of the manufacturer.

For expression, cells were grown in baffled flasks with 800 mL LB medium supplemented with 50 µg mL^{-1} ampicillin and arabinose concentrations of 0,002% (FMO-E and FMO-G) or 0,2% (FMO-F), shaking at 200 rpm at 24°C for 32 h, based on the optimal expression conditions determined before (Riebel *et al.* 2013). Cell extracts were prepared by sonication with subsequent centrifugation to obtain the cleared cell extracts. Purification of the enzymes from the cleared cell extracts was done with *Strep*-Tactin® Sepharose from IBA GmbH, following the recommendations of the manufacturer. As minor modification, 10 µM FAD, 1,0 mM DTT and 10% glycerol were added to the buffer.

Determination of kinetic parameters of FMO-E

Activities of purified FMO-E were determined spectrophotometrically by monitoring increase or decrease of NADPH or NADH respectively in time at 340 nm. The reaction mixture (1,0 mL) typically contained 50 mM Tris/Cl, pH 7,5 (10% glycerol, 1 mM DTT, 1 mM EDTA, 10 µM FAD), 150 µM coenzyme NADPH or NADH, 50 µM ketone substrate,

5,0 % dioxane and 0,1 µM enzyme at 25 °C. Kinetic parameters were obtained by fitting the data as described previously (Torres Pazmiño *et al.* 2008).

Stopped-flow experiments

Both, the reductive and oxidative half-reactions of FMO-E were analyzed using the Applied Photophysics SX17MV stopped-flow instrument equipped with a photodiode array detector. All experiments described were performed at 25 °C in 50 mM Tris/Cl (10% glycerol, 1,0 mM DTT, 1,0 mM EDTA), pH 7,5. All concentrations stated are those in the reaction chamber after mixing. Anaerobic conditions were achieved by flushing the system and solutions with N_2 and removing traces of oxygen upon addition of 1,0 mM putrescine and a catalytic amount of putrescine oxidase. The rate of reduction of FMO-E by NADPH was assessed under anaerobic conditions by adding equimolar amounts of NADPH. The reoxidation of the reduced flavin was achieved by mixing NADPH-reduced FMO-E with aerated buffer and following the spectral changes of the flavin cofactor in time. The obtained spectra were analyzed by means of numerical integration methods using Pro-K (Applied Photophysics Ltd.), yielding the observed rate constants.

Bioconversions with cleared cell extracts

Conversions were performed as described before (Riebel *et al.* 2013). For determining the exact concentration of each enzyme in the respective extract, a recently developed method was used which relies on the decrease in absorbance at 450 nm upon NADPH-mediated reduction of the flavin cofactor (Riebel *et al.* 2012).

Whole-cell conversions

Bioconversion of a set of 14 target compounds was performed by using cells expressing FMO-E, FMO-F and FMO-G employing the above-mentioned expression conditions. Cells were grown up to OD_{590} = 0,2–0,6 at 37 °C. The cultures were then cooled to 24 °C and L-arabinose was added as 10% w/v solution in water to the appropriate final concentration. Cyclodextrin (4,0 mM) was supplemented as cell membrane transfer agent. This mixture was thoroughly shaken and then divided into aliquots (1,0 mL) in 24-

well plates. Substrates were added as solution in 1,4-dioxane (5,0 mM final concentration), the plates were sealed with adhesive film (EasySEAL, Greiner Bio One) and incubated in an orbital shaker (Infors HT Multitron 2 Standard) at 24 °C for 24 h. Analytical samples were prepared by extraction of 0,5 mL of biotransformation culture with 1,0 mL Ethylacetate (supplemented with 1,0 mM methyl benzoate as internal standard) after centrifugal separation of the cell mass (approx. 15 kRCF, 1 min, rt). All transformations were carried out as technical triplicates; conversion and selectivity are reported as arithmetic mean values. Absolute configuration or optical rotation values of lactone products were assigned on the basis of published reference biotransformations for substrates **4-9** (Rudroff *et al.* 2007), for substrate **11** (Fink *et al.* 2012), and for substrates **1 & 14-16** (Mihovilovic *et al.* 2008**)**.

In summary there are no small problems.

Problems that appear small are large problems that are not understood.

Santiago Ramòn y. Cajal

CHAPTER 6

SUMMARY & FUTURE PERSPECTIVES IN YELLOW

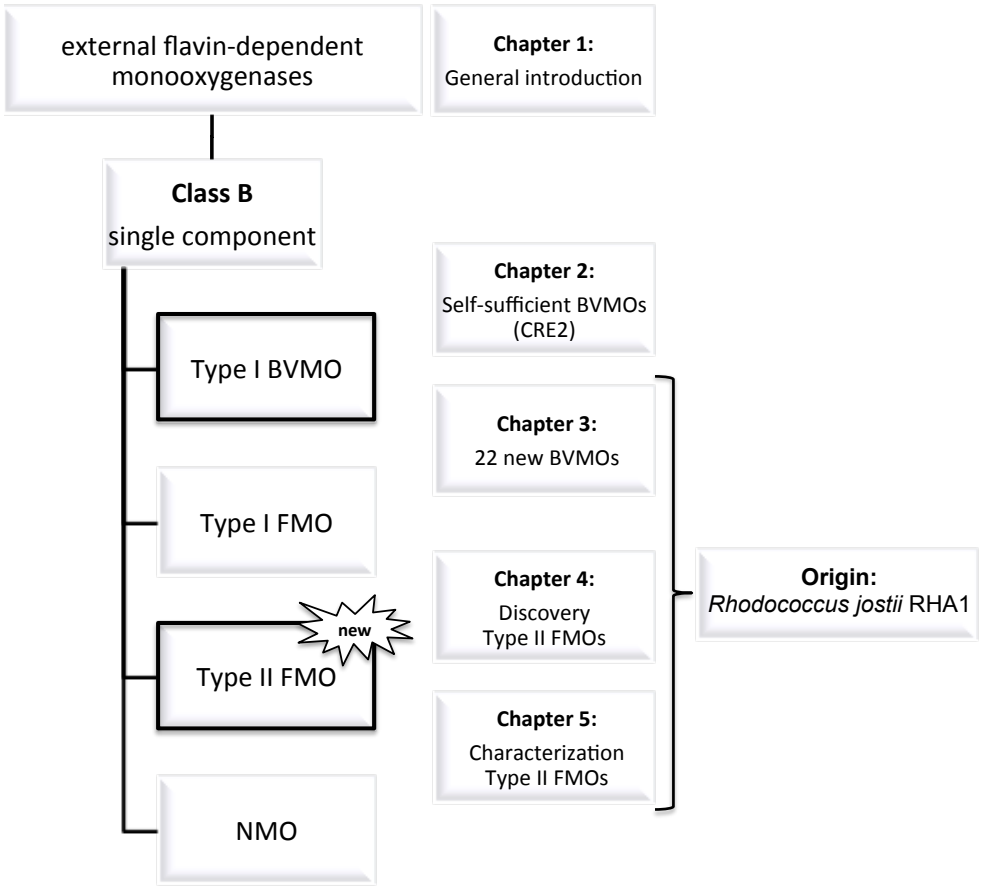


Figure 6-1: graphical outline of the thesis

6.1. Introduction and general overview

Flavoenzymes have mastered a huge range of reactions. Therefore, many flavin-dependent enzymes show great promise as biocatalysts for industrial applications. This increases the need to make these enzymes applicable and easy to handle. Besides the limited availability of flavoenzymes, their application is often difficult due to cofactor dependency, especially when a flavoenzyme is strictly dependent on the expensive NADPH coenzyme.

In this thesis the above-mentioned topics have been addressed (Figure 6-1). To increase the amount of available applicable flavoenzymes, a large number of flavoprotein monooxygenases from the soil organism *Rhodococcus jostii* RHA1 have been produced and studied (**chapter 3, 4, 5**). **Chapter 2** offers a general solution to the nicotinamide coenzyme dependency of flavoprotein monooxygenases by a systematic optimization of a coenzyme regeneration system (CRE2).

Chapter 4 and **5** deal with the two issues mentioned above in a special way. With the discovery of Type II FMOs, a new subclass of class B flavoprotein monooxygenases has been found that is able to utilize both nicotinamide coenzymes, NADPH and NADH. Members of this subclass of monooxygenases were found to catalyze Baeyer-Villiger oxidations. Therefore, these newly discovered biocatalysts make it possible to use NADH instead of NADPH for enzyme-catalyzed Baeyer-Villiger oxidations.

6.2. Cofactor regeneration by fusion (chapter 2)

Flavoprotein monooxygenases usually utilize nicotinamide coenzymes. The supply of these expensive cofactors is often a major bottleneck for the industrial use of these biocatalysts. By fusing a Baeyer-Villiger monooxygenase (BVMO) with an NAD(P)H-regenerating enzyme, a self-sufficient and easy-to-handle system (CRE = coenzyme regenerating enzyme) was developed earlier in our laboratory (Torres Pazmiño *et al*. 2008). For this approach phosphite dehydrogenase from *Pseudomonas stutzeri* (PTDH) had been chosen as fusion partner as it has been engineered by Woodyer *et al.* towards a

relaxed nicotinamide cofactor specificity (Woodyer *et al.* 2003). In addition, no interference of the substrates and products of both enzymes, the monooxygenase and PTDH, have been detected. Furthermore, due to the favorable thermodynamic equilibrium constant, the oxidation of the extremely cheap sacrificial cosubstrate phosphite is an almost irreversible process (Woodyer *et al.* 2003). Experiments confirmed that the covalently fused PTDH-BVMOs efficiently work well in crude-cell extracts or purified form. As the fusion of the two biocatalysts had no significant influence on the performance of each entity it proves to be an efficient approach.

Fusing PTDH at the N-terminus of the targeted BVMO was found to be beneficial for boosting overexpression of the bifunctional biocatalyst (Torres Pazmiño *et al*. 2008). For the development of an improved second-generation expression vector for the production of PTDH-BVMOs (pCRE2), a PTDH-encoding gene optimized for expression in *E. coli* was used. Moreover, PTDH was equipped with an N-terminal polyhistidine-tag. In this way the PTDH can be employed as NAD(P)H regeneration system, while it also facilitates expression and purification of such self-sufficient bifunctional biocatalysts. Furthermore, the optimized gene encoding for PTDH contains 18 mutations that were known to improve its stability (Woodyer *et al*. 2005, Johannes *et al*. 2006, Woodyer *et al.* 2006). This 18-fold PTDH mutant shows much higher thermostability and solvent resistance compared to the wild-type PTDH. With the original pCRE construct, the fused PTDH was found to be relatively labile. By using the 18-fold PTDH mutant, the PTDH fusion enzyme was shown to be superior to all studied monooxygenase fusion partners.

We also tried to further improve the pCRE2 expression vector by changing the linker length and composition between PTDH and the targeted enzyme. In the original PTDH-BVMO fusion proteins, we observed a slightly negative effect upon flavin binding of the BVMOs. However, we could not identify a better-suited linker concerning the binding of the FAD cofactor. The effects of various linker versions were only marginal. Based on these findings, we continued the research using the pCRE2 expression vector with the original linker for fusing PTDH to target proteins.

In **chapter 2**, we also describe the evaluation of the pCRE2 expression system by producing six PTDH-BVMOs (Torres Pazmiño *et al.* 2009). Besides BVMOs, we could also show the applicability of the pCRE2 system for other NADPH-dependent enzymes. Recently a bacterial flavin-containing monooxygenase (FMO), mFMO from *Methylophaga* sp. strain SK1, was shown to be successfully produced by using the pCRE2 vector (Rioz-Martínez *et al.* 2011). Additionally, the gene encoding for the P450 monooxygenase P450BM3 from *Bacillus megaterium* was cloned into pCRE2 and the PTDH-P450BM3 could efficiently be expressed and purified (data not published). The PTDH-P450BM3 fusion protein was shown to be active as self-sufficient monooxygenase. This shows that the pCRE2 system also works for large, multidomain enzymes. For the expression of the enzymes discussed in this thesis (**chapter 3 & 4**), the pCRE2 expression vector also played an important role. By using this fusion-expression system it was possible to express all targeted proteins in significant amounts to test their substrate acceptance. When using the pBAD*N* expression vector only around half of them would have been available for further characterization.

The approach of producing recombinant redox biocatalysts fused to PTDH will facilitate future studies of NAD(P)H dependent biocatalysts. By equipping target enzymes with PTDH, the resulting fusion enzyme is not only able to regenerate NAD(P)H, but it also allows the use of a colorimetric assay for quantifying the enzyme activity (**chapter 3**). In addition, in all cases tested so far, pCRE2 improved the expression levels when compared with expressing the native protein without a fusion partner.

As neither the composition nor the length of the linker has a drastic effect on the performance of the fused PTDH-BVMOs, the design of a linker with a protease recognition site may be considered in future studies. This will allow the removal of the PTDH moiety for studies on the native target enzyme. A suited protease seems to be the TEV-protease from tobacco etch virus as it leaves only an additional glycine at the N-terminus of the cleaved protein. Therefore it allows production of a nearly native form of the target enzyme. A potential problem of this protease is that it recognizes a protein sequence mainly consisting of amino acid residues strongly avoided in natural occurring linker

sequences: **EN**LYF**F**QG (the un-preferred residues according to Argos are marked in bold (Argos 1990)). Yet, as we could show that the linker sequence does not have a huge effect on the catalytic performance of the fusion protein, it might be of interest to implement this sequence in the linker of the next generation pCRE: pCRE3.

6.3. The oxidative treasures of *Rhodococcus jostii* RHA1

In silico screening of the genome of *Rhodococcus jostii* RHA1 revealed that this organism harbors a huge collection of putative oxidative enzymes. By using sequence motifs and sequence homology searches, we could identify 40 (putative) flavoprotein monooxygenase-encoding genes. When compared with other bacterial genome sequences, this is an exceptionally large number of flavoprotein monooxygenases for one organism. Except for these monooxygenases, *R. jostii* RHA1 also harbors many other oxidative enzymes (McLeod *et al.* 2006).

To explore the biocatalytic potential of the rhodococcal monooxygenases, all putative BVMO- and FMO/NMO-encoding genes were efficiently cloned and expressed in *E. coli* using the pBAD*N* and the pCRE2 expression vector. The biocatalytic scope of these enzymes is described in **chapter 3, 4** and **5**. Except for these class B flavoprotein monooxygenases, also Type II BVMO- and LadA-encoding genes have been cloned. Unfortunately, characterization attempts of these putative monooxygenases failed, mainly due to expression problems. This shows that the pCRE2 is not always of help. For these problematic cases, homologous overexpression in *R. jostii* RHA1 could be a solution (Montersino 2012).

6.3.1. Type I BVMOs from *Rhodococcus jostii* RHA1 (chapter 3)

The work described in **chapter 3** confirms that *Rhodococcus jostii* RHA1 harbors 23 genes that encode for Type I BVMOs. Two of them were found to differ in only one amino acid upon which it was decided to exclude one from our studies. On the other hand we cured one gene that suffered from a one-nucleotide-deletion in the genomic sequence and

made it available again. By expressing and subsequent substrate screening of 22 Type I BVMOs we increased the number of characterized BVMOs drastically. Sequence alignment of the BVMOs together with some of the best characterized BVMOs (Figure 3-2) revealed that the rhodococcal BVMOs represent many different clades of the Type I BVMO family.

New BVMO identifying motif

By a multiple sequence alignment of Type I BVMO sequences we could see that the previously described Type I BVMO typifying sequence motif FxGxxxHxxxW[P/D] (Fraaije *et al.* 2002) is not always strictly conserved. In a number of sequences one or two variations were observed. As this BVMO motif is located in a loop region that allows spatial domain-domain rearrangements during the catalytic cycle, it consists of residues that are not directly involved in catalysis (Orru *et al.* 2011). Using sequences from already characterized Type I BVMOs together with the 23 rhodococcal sequences, we identified another conserved motif, located between the N-terminal GxGxxG motif and the known BVMO motif: [A/G]GxWxxxx[F/Y]P[G/M]xxxD.

The conservation of several residues in this motif is in full agreement with the fact that they interact with the flavin cofactor and/or play a crucial role in catalysis, *e.g.* Asp66 in PAMO is involved in the direct interaction with the $NADP^+$ coenzyme and mutating this residue resulted in inactivation of the enzyme (Orru *et al.* 2011). As this motif entails more conserved residues when compared with the other BVMO-typifying motif, this newly recognized motif appears to be more discriminative for identifying Type I BVMOs.

Unraveling the substrate profile

We have determined the substrate profiles of all studied rhodococcal BVMOs. For such large scale substrate profiling, several technical procedures have been developed. To allow a fast and easy high throughput assay for BVMO activities in cell extracts we developed two new methods. One method allows determination of the concentration of recombinant BVMO in cell extract. The second method involves a novel approach to assay for BVMO activity in microtiterplates.

Concentration determination via flavin reduction with NADPH

As all Type I BVMOs contain an FAD cofactor that is specifically reduced by NADPH, we developed a simple method for accurate BVMO concentration determination. For this method we were inspired by the well-known method to quantify the amount of P450 monooxygenases, which uses the specific heme absorption features of reduced P450 monooxygenases in complex with carbon monooxide. In our method, we measure the absorbance spectrum of the cell extract at 440 nm before and after the addition of NADPH (Figure 3-3). With an extinction coefficient that reflects the absorbance difference between oxidized and reduced BVMO, the amount of NADPH-reactive flavin cofactor in the expressed BVMO can be determined. As only one flavin is bound per BVMO, the flavin concentration is a good measure for the concentration of active BVMO. The method is very selective for the recombinantly (over)expressed BVMO as only a minority of the native proteins in *E. coli* contains a flavin cofactor and most of them do not react with NADPH. Validation of the method revealed that it can be applied to cell extracts and purified enzymes. The method should also be applicable to purified NADH dependent flavoenzymes.

Substrate screening: phosphate assay

The original idea for the substrate profiling studies was to apply the often-used photometric method which monitors the NADPH consumption at 340 nm during the reaction. Using the initial depletion rates in combination with the enzyme concentration, the observed rates can be determined. However, this method turned out not to be a suitable one. The following factors prevented the use of this widely applied assay procedure:

- In order to normalize the amount of each BVMO in each assay different amounts of cell extract had to be used, depending on the expression level of the respective BVMO. Hence the background absorption shifted in every measurement, which made it difficult to compare the results with each other.

- In order not to exceed the capacity of the instrument, only relatively low concentrations of NADPH could be used.
- Many targeted compounds absorb light in the same wavelength regime as NADPH (300–390 nm).
- The initial rate of NADPH consumption does not reflect the real potential of a biocatalyst in terms of efficiency, as this depends on factors like stability and inhibitory effects. Only a longer reaction period allows probing these factors.

To overcome these problems we developed an indirect assay by coupling the oxidation reaction of the BVMO with the NADPH regeneration of the phosphite dehydrogenase (PTDH). The assay is based on the colorimetric detection of phosphate which is formed by PTDH while regenerating NADPH. As it is an indirect measurement, all the above-mentioned issues do not play a role. For example, this set-up allows monitoring conversions over long periods of incubation.

Using the developed assay we were not only able to give a substrate profile overview of the newly found rhodococcal BVMOs, but also to verify and supplement the list of confirmed substrates for a number of known BVMOs: PAMO, STMO, CPMO, CPDMO (not published), HAPMO and CHMO. The generic assay proved to be reliable and robust, thereby representing a new tool for substrate profiling of nicotinamide coenzyme dependent enzymes. With the relaxed specificity of PTDH towards the nicotinamide coenzyme, it can be used for any enzyme that requires NAD(P)H.

Substrate scope

The results in **chapter 3** disclose a broad substrate scope for the studied set of 22 BVMOs. Upon screening 39 potential substrates (Table 3-3), at least one active enzyme could be found for each substrate. However, most of the *Rhodococcus* BVMOs are not so efficient biocatalysts: eight BVMOs (BVMO #1, #3, #6, #7, #8, #13, #16, and #17) did not show significant activity for any of the tested compounds. Nevertheless, BVMO #3, #7, and #17 produced indigo blue when expressed in *E. coli* (Table 3-1), suggesting that they are active, even though not very efficient on any of our test panel compounds. Most of the

BVMOs active on the tested substrates displayed a rather limited substrate scope. Only five BVMOs (BVMO #4, #15, #20, 2#1, 2#4) were shown to be active on ≥10 tested compounds. BVMO #4 and BVMO #24 displayed the broadest substrate scope of all tested BVMOs. In a phylogenetic tree based on the core active site residues (Figure 3-2), the BVMOs with a relaxed substrate acceptance profile were found to be closely related to the clades that include the known prototype BVMOs (CHMO, CPMO, PAMO and STMO). The less active rhodococcal BVMOs cluster with groups that include other known BVMOs (HAPMO, MoxY, EtaA, AlmA) which are known for their limited substrate scope. This indicates that it may be more productive to focus on homologs of the typical prototype BVMOs when looking for novel and biocatalytically relevant BVMOs.

To verify the results of the indirect assay and to gain more information on the biocatalytic potential, we evaluated a selected number of BVMOs as biocatalysts with typical BVMO substrates and prochiral sulfides. The results confirmed the outcomes of the substrate screening assay and revealed that the sulfides were efficiently converted with excellent enantioselectivities. Again, unique capabilities of BVMO #24 were observed, as it displays opposite enantioselectivities when compared with the other tested BVMOs. For further studies, the most exciting candidates are BVMO #4 and BVMO #24. They show the most interesting activities with a broad substrate scope and, in case of BVMO #24, atypical enantioselective behavior.

6.3.2. FMOs and NMOs from *Rhodococcus jostii* RHA1 (chapter 4 & 5)

Except for the genes encoding for Type I BVMOs, the *Rhodococcus jostii* RHA1 genome contains ten additional genes that encode for proteins with two Rossmann-fold motifs, indicative for class B flavoprotein monooxygenases. Most of these predicted flavoproteins resembled FMOs as they contain the FMO motif. Already the finding of 23 Type II BVMOs in the genome of one organism was spectacular (the average is one or two BVMO-encoding genes per microbial genome). Finding even more class B flavoprotein monooxygenases is quite remarkable, as FMOs are even more rare in bacteria when

compared with BVMOs (de Gonzalo *et al.* 2010). All ten genes have been cloned, expressed and tested for activity as previously described for the BVMOs.

Three of the studied enzymes were found not to belong to the FMO family. One of them is an NMO (NMO-H, **chapter 4**), while another gene appears to encode for a tryptophan monooxygenases. The third enzyme is a putrescine oxidase homolog (unpublished results). The remaining seven protein sequences share the FMO-typifying sequence motif, with only one or two amino acid substitutions (Table 4-1). They cluster together in one branch of a phylogenetic tree of class B flavoprotein monooxygenases. They form a distinct group of sequences different from known FMOs (Figure 4-1). As a consequence, we classified them as Type II FMOs, while naming the known and well-established FMOs Type I FMOs.

One interesting fact and clear justification for the introduction of a new class of subclass B flavoprotein monooxygenases was the discovery that the Type II FMOs show no coenzyme preference – they accept both, NADH and NADPH with similar affinity as reduction equivalents. This is a unique and new feature of class B flavoprotein monooxygenases. According to the expected substrate scope for FMOs, the new Type II FMOs were first tested for activity towards reactions typically catalyzed by FMOs: the oxidation of sulfides and amines. We successfully adapted a colorimetric assay for the conversion of indole and indole-derivatives described in the work of Rioz-Martinez *et al*. (Rioz-Martínez *et al.* 2011) for the use with cell extracts (data not published). Surprisingly, none of the tested indoles (in total 20 derivatives) and amines was converted by any of the new Type II FMOs. This indicates that the natural substrates for these enzymes might be completely different when compared with the substrate scope of Type I FMOs. Sulfoxidation reactions were only catalyzed with low rates. Three of the Type II FMOs (FMO-E, -F, and -G) showed higher conversions in the sulfoxidation of thioanisole then the other four. These three enzymes were, in contrast to the other Type II FMOs, also able to convert larger and bulkier sulfoxides, albeit with very low activity. The most remarkable feature of these three enzymes was the fact that they were also able to convert ketones, yielding the products of a Baeyer-Villiger oxidation. This reaction is rarely catalyzed by

FMOs and suggests that this subgroup of Type II FMOs (FMO-E, -F, and -G) has evolved towards *bona fide* Baeyer-Villiger monooxygenases. Especially in combination with the relaxed coenzyme preference, they represent a very interesting group of enzymes for biocatalysis. FMO-E, -F, and -G, are closely related in sequence and share an *N*-terminal extension of about 160 residues. Using such N-terminal sequence, hundreds of homologous putative Type II FMOs can be identified in the databases, which are mainly from fungi. It will be interesting to see whether the Baeyer-Villiger monooxygenase activity and the relaxed coenzyme specificity can also be found in all these homologs. If so, they might represent interesting alternative biocatalysts compared to known BVMOs. Additionally, it might well be that the observed BVMO activity is just a promiscuous side activity and the original function for these Type II FMOs still has to be explored.

In **chapter 5** we set out to investigate the three special Type II FMOs, FMO-E, -F, and -G in more detail. Besides analyzing the enantioselectivity towards thioanisole and bicycloheptenone (Table 4-3 and Table 4-2), the three enzymes were used to convert a set of cyclohexanone and cyclobutanone derivatives. While the cyclohexanones where not converted, the cyclobutanones were oxidized into the corresponding lactones with modest to good enantioselectivity (Table 5-3 and Table 5-2). In the conversion of racemic fused cyclobutanones (Table 5-2) they disclosed new and excellent stereoselectivities, better than ever observed for any known BVMO in the case of norcamphor.

For a more detailed characterization we purified FMO-E using a strep-tag. The enzyme displays a typical flavin absorbance spectrum (Figure 5-2). In steady-state kinetic experiments FMO-E showed typical Michaelis-Menten behavior with the two coenzymes NADH and NADPH and the Baeyer-Villiger substrate bicycloheptenone. High k_{cat} values confirm that it is a competent enzyme: the rate of catalysis is in the same range as has been found for other class B flavoprotein monooxygenases. FMO-E also shows high affinities towards both nicotinamide coenzymes with both K_M values in the low µM range.

To establish the reaction mechanism, FMO-E was studied by using the stopped-flow technique (Figure 5-3). Pre-steady state kinetic analysis revealed features known for class B flavoprotein monooxygenases: (i) oxidized FMO-E is rapidly reduced by a reduced

coenzyme without the need of substrate binding, (ii) the reduced FMO-E rapidly reacts with molecular oxygen, and (3) FMO-E is able to form a stable peroxyflavin intermediate. With the identification of the peroxyflavin intermediate, which is considered to be the crucial oxygenating species, we could show that these enzymes operate similar to Type I BVMOs.

The next step was a careful inspection of the predicted active site residues to find out the differences between the Type II FMOs when compared with "real" Type I BVMOs and Type I FMOs. For BVMOs and FMOs it is quite well known which residues are important for catalysis. As no structural data is available for the Type II FMOs, the analysis was performed on a homology model. It became obvious that all three enzymes employ different strategies to catalyse oxygenation reactions: even though the Type II FMOs have conserved residues (His222 & Arg563 in FMO-E) on similar positions close to the isoalloxazine ring as the other two class B enzymes, they form a different active site architecture (Figure 5-4). We have targeted these two residues in a site-directed mutagenesis strategy by alanine replacements. We observed that mutating these residues has an effect on cofactor binding. Purification experiments only yielded the inactive apo form without bound FAD and it was impossible to restore the activity by adding the flavin cofactor afterwards. Biocatalysis performed with the mutants as cell extracts showed that the mutant enzymes show no BVMO activity. These results imply that these two residues obviously play an important role in binding of the flavin cofactor.

All these data confirm that the newly discovered subclass of Type II FMOs has very interesting catalytic capabilities, especially in combination with the preference for NADH. This not only sets them apart from Type I BVMOs and FMOs, but makes them valuable biocatalysts on their own. Further structural, kinetic and mutagenesis studies will reveal the details on their functioning. It will also be interesting to find out what the native substrates for these biocatalysts are. With many homologs in fungal genomes, this class of enzymes represents a promising and exciting new player in the world of flavoenzymology.

I believe in intuition and inspiration.

Imagination is more important than knowledge.

For knowledge is limited, whereas imagination embraces the entire world,

stimulating progress, giving birth to evolution.

It is, strictly speaking, a real factor in scientific research.

Albert Einstein, Cosmic religion 1931

APPENDIX

GC conditions

Tested compounds in the phosphite screening assay

Multiple sequence alignment of BVMO protein sequences

Sequence alignment of FMO #E, #F, #G

All protein sequences of the cloned *Rhodococcus jostii* RHA1 genes

Purification of Type II FMO E

GC conditions

The following columns were used for the determination of conversions and enantiomeric excesses of the sulfoxides by GC:

A: Alltech GT-A (30 m x 0,25 mm x 0,25 µm),

B: Hewlett Packard HP-1 (30m x 0,32 mm x 0,25µm, 12,2 psi N2),

C: Chirasil Dex CB (30 m x 0,25 mm x 0,25 µm, 12 psi N2).

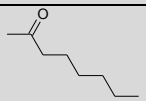
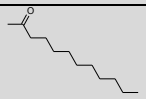
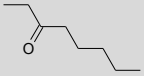
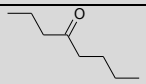
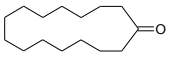
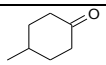
Table A-1: GC-conditions for the determination of the conversions and enantioselectivities.

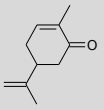
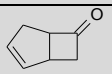
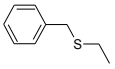
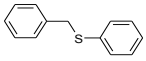
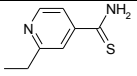
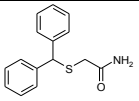
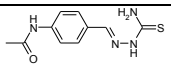
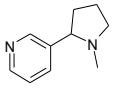
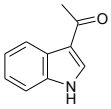
Substrate	Program	Column	t_R (min) substrate	t_R (min) products
thioanisole	70/5/10/200/2	B	11,4	14.9 (sulfoxide) 15.6 (sulfone)
thioanisole	40/0/10/160/8	A	7,5	14.3 (R); 16.7 (S)
p-tolyl methyl sulfide	70/5/10/200/2	B	11,5	15.1(sulfoxide) 15.9 (sulfone)
p-tolyl methyl sulfide	40/0/10/160/8	A	8,3	15.6 (R); 16.5 (S)
benzyl ethyl sulfide	70/5/10/200/2	B	12,5	14.7 (sulfoxide) 16.9 (sulfone)
benzyl ethyl sulfide	40/0/10/160/8	A	8,8	14.4 (R); 15.5 (S) 19.0 (sulfone)
benzyl phenyl sulfide	100/5/5/200/5	B	21,7	27.3 (sulfoxide)
2-indanone	70/10/3/180/5	B	22,3	32,7
cyclopentadecanone	70/10/3/180/5	B	46,5	50,4
phenylacetone	70/7/5/120/2	B	13,5	14.9 (ester) 10.3 (alcohol)
bicyclohept-2-en-6-one	70/5/5/150/0	B	6,6	13.3 (abnormal; AB) 13.4 (normal; N)
bicyclohept-2-en-6-one	130℃ isotherrm	C	9.7 and 9.9	(AB: 1*R*,5*S*): 17.6 (N: 1*R*,5*S*): 18.0 (AB: 1*S*,5*R*): 18.3 (N: 1*S*,5*R*): 18.5

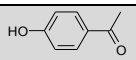
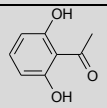
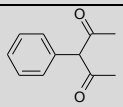
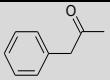
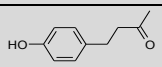
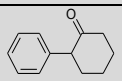
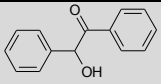
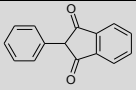
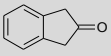
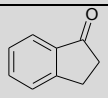
[a] **Program: initial T (°C)/ time (min)/ slope (°C/min)/T (°C)/ time (min).**

Tested compounds in the phosphite screening assay

Table A-2: Name, CAS number and structural formular of the used assay substrates.

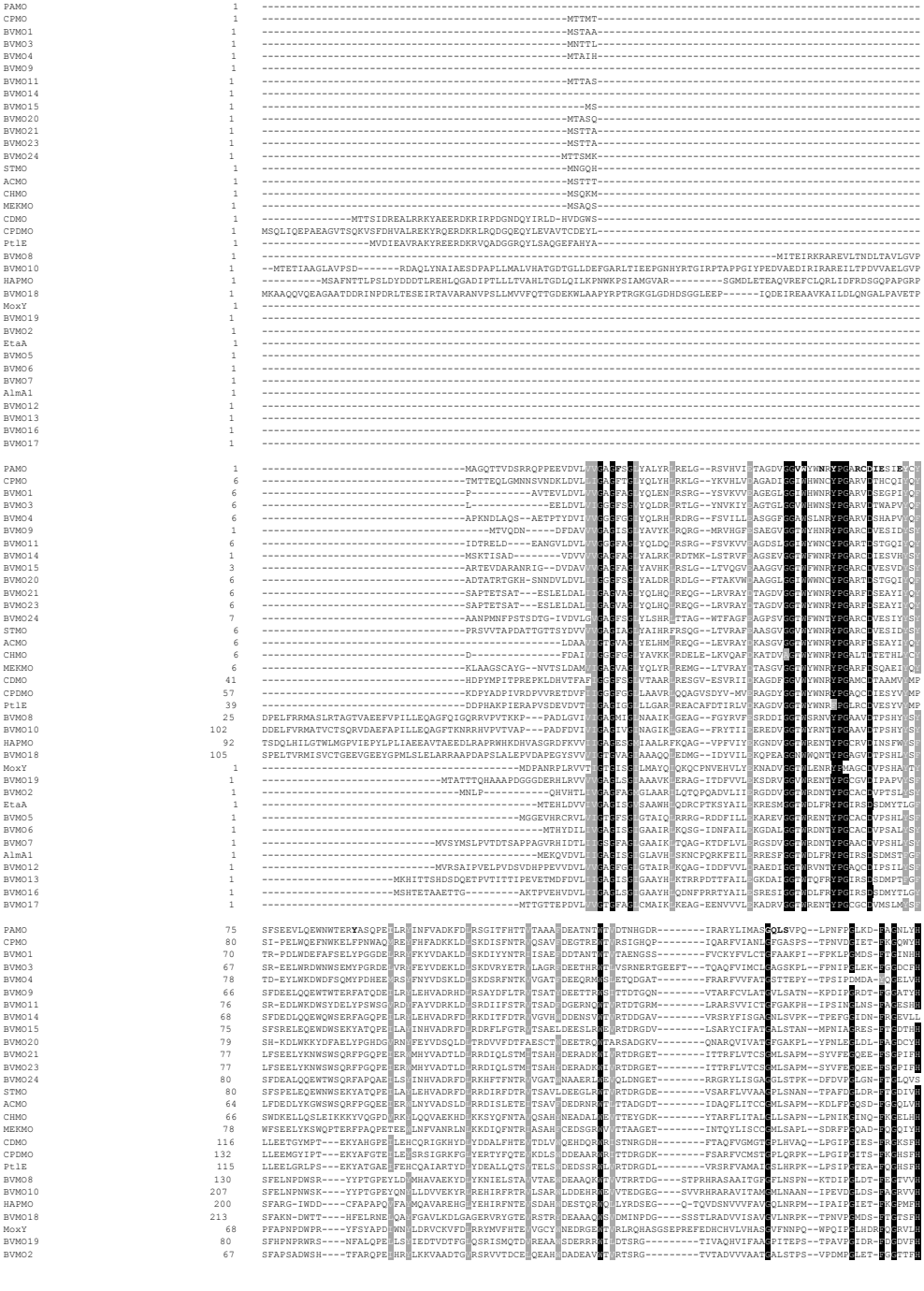
Compound name	CAS number	structure
Acetone	67-64-1	
2-Butanone	78-93-3	
2-Octanone	111-13-7	
2-Dodecanone	6175-49-1	
3-Octanone	106-68-3	
4-Octanone	589-63-9	
Methylvinylketone	78-94-4	
3-Methyl-2,4-pentanedione	815-57-6	
Cyclobutanone	1191-95-3	
Cyclopentanone	120-92-3	
Cyclohexanone	108-94-1	
Cyclopentadecanone	502-72-7	
Oxocyclohexanecarbonitrile	4513-77-3	
4-Methylcyclohexanone	589-92-4	

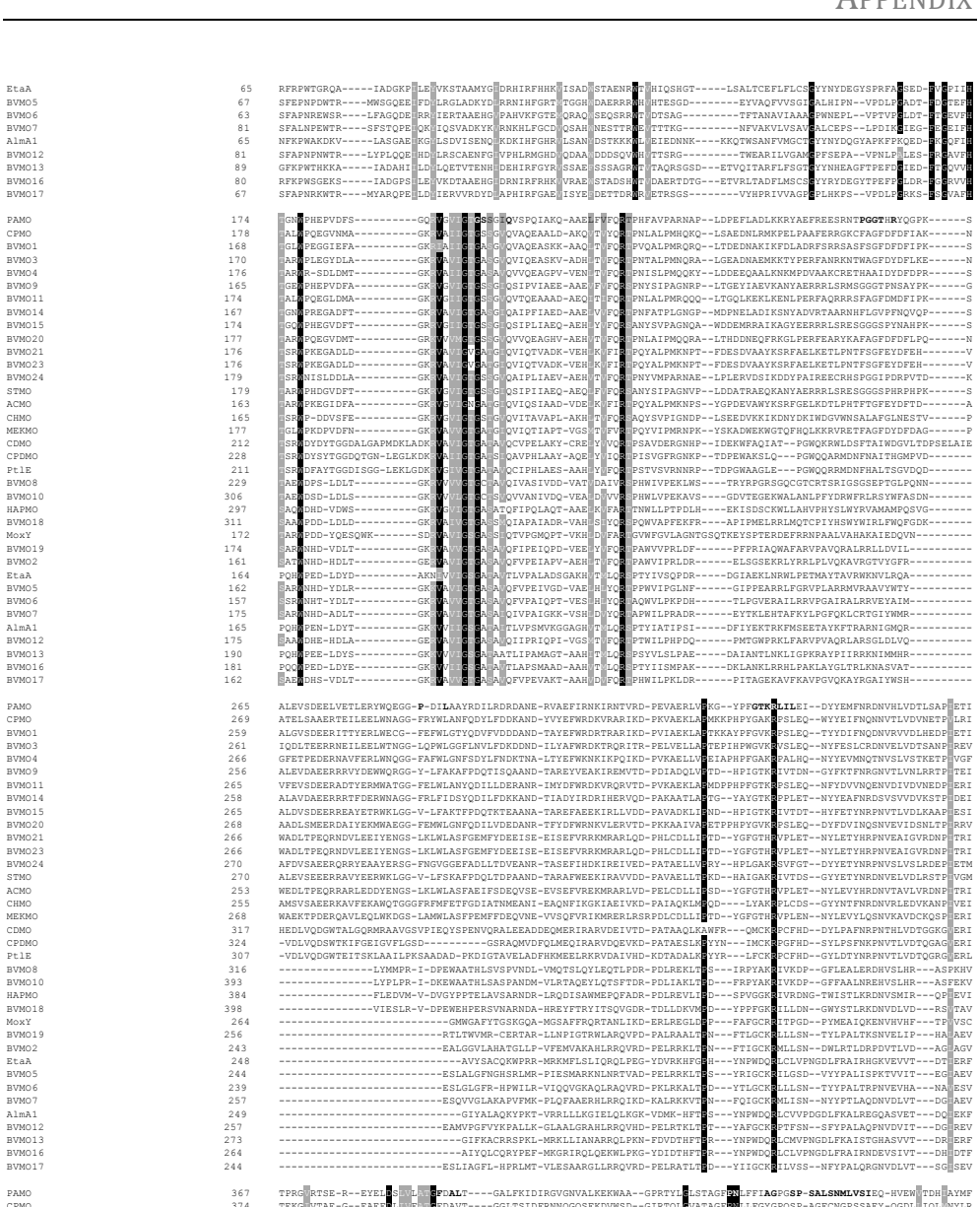
2-Propylcyclohexanone	94-65-5	
Dihydrocarvone	7764 50 3	
Cyclopropylmethylketone	765-43-5	
Norcamphor	497-38-1	
Bicycloheptenone	13173-09-6	
Thioanisole	100-68-5	
Benzylethyl sulfide	6263-62-3	
Benzylphenyl sulfide	831-91-4	
Ethionamide	536-33-4	
Modafinil precursor	68524-30-1	
Thiacetazone	104-06-3	
Nicotin	54-11-5	
Indole	120-72-9	
3-Acetylindole	703-80-0	
Acetophenone	98-86-2	

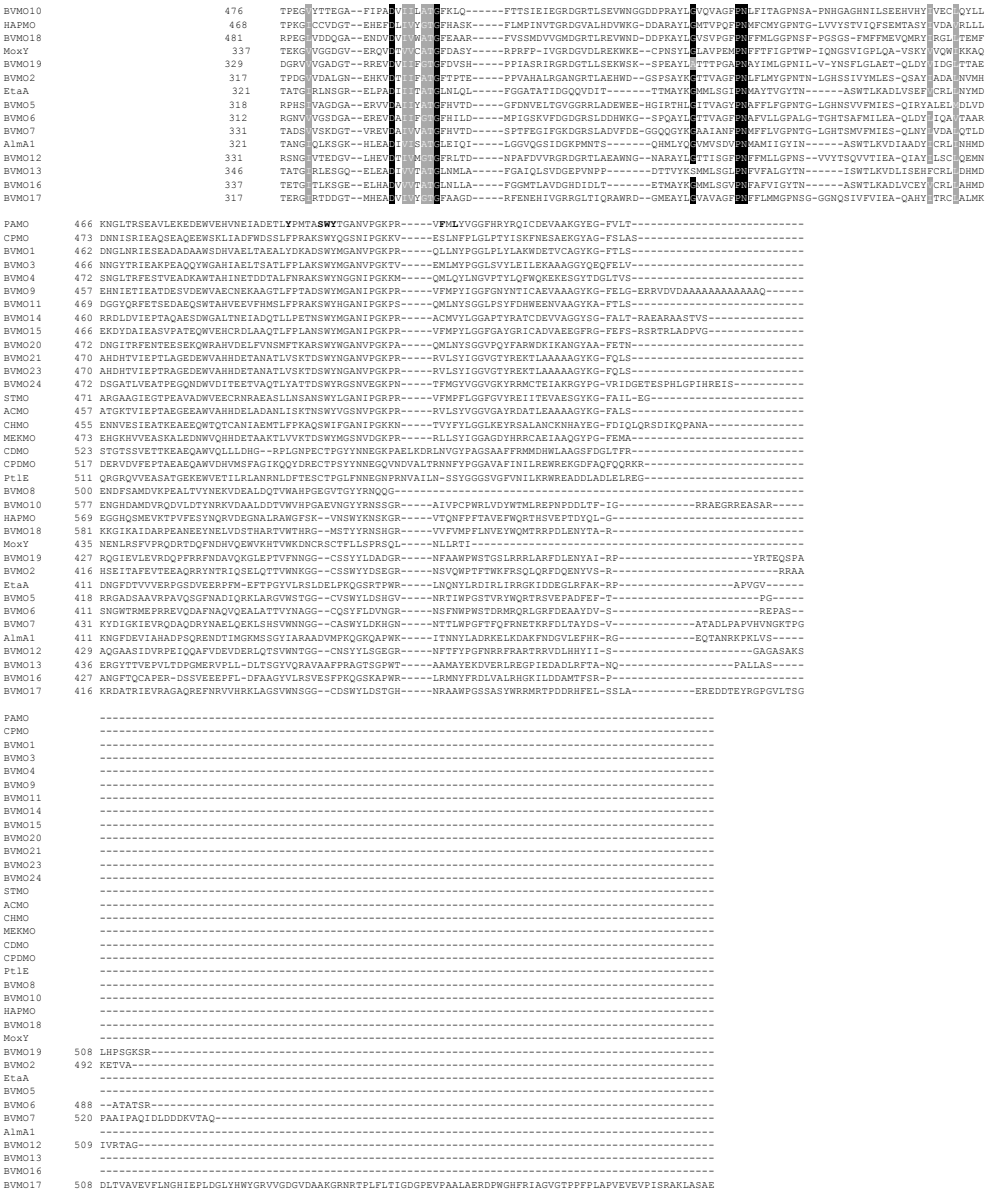
4-Hydroxyacetophenone	99-93-4	
2,6-Dihydroacetophenone	699-83-2	
3-Phenylpentane-2,4-dione	5910-25-8	
Phenylacetone	103-79-7	
Raspberry ketone	5471-51-2	
2-Phenylcyclohexanone	1444-65-1	
Benzoin	119-53-9	
Phenendione	83-12-5	
2-Indanone	615-13-4	
1-Indanone	83-33-0	

The shading indicates different substrate classes.

Multiple sequence alignment of BVMO protein sequences







Sequence alignment of FMO #E, #F, #G

```
f               MTQTVQPAAVQTSLTPQERVDLWLASFESALAARDVDRAAGMFAVDSFWRDLVAFTWNLK 60
g               MTATLDAPVDTNIPQPGDIARRWLAGFGATLERGDARGAAQHFLVDGWWRDLLSFTWDLH 60
e               MTTTFSDTDLRTD------AQAWLDGFSRFLAAELAP--TAVFAPQAYWRDVLAFTGDLR 52
                ** *.. .   .       .  ** .*   *    .   :  *  :.:***:::** :*:

f               TVEGRDAVAAMLHARLDDTDPVNFRTT-----ETPDEADG--VTSAWIEFETATGRGKGH 113
g               TTHGRADIESRLADSVPVHEPRHLVLSPAHPAEAVADPEGD-WIQAFFTFETTLARSRGF 119
e               TFS--DEIPAELLRRQELTKATNIRIAEDRTPPRLVERAGIPCLEVIFEFDTLAGSAVGV 110
                *      : : *       .. ::  :         :  *    .. : *:*  . . *


f               LRLKGDE-----AWTFLTTMQELKGHEERRGRNRVKGAVHGSGGDTLSWAEKREIEEREL 168
g               VRLRRDDDGEWRAWTLISAMEEIKGHEEKKGHRRVQGTNHGAHRGKINWLDRR-TAKGEF 178
e               ARLVDVPERGLLVRSLFTTLDQLADHPERTGEHRPVGQADSSKFGGPNWLDRR-IAAQAY 169
                 **         . ::::::::: .* *: *..*  *  ..:  .  .* ::*

f               GYTRQPYVLVIGGGQGGIALGARLRQLGVPAIVVDKNERPGDQWRNRYKSLCLHDPVWYD 228
g               -ETEQPAVVIVGAGQGGLALAARLGQLGVDTLLVERNDRIGDSWRKRYHSLVLHDPVWYD 237
e               -ENRDPDVLIVGGGQSGLTLAARLGQLDVDALVVDTHARPGDNWRTRYHALTLHNAVWLN 228
                  ..:* *:::*.**.*::*.*** **.* :::*: : * **.**.**::* **:.** :

f               HLPYMPFPDNWPVFAPKDKIGDWLEMYTKVMEIPYWSSTTCTSATFDDETKEWTVVLDRD 288
g               HLPYLNFPDHWPVFTPKDKLANWFEFYADAMELNVWTGTEFTGGSYDDATGEWTVTVARD 297
e               DLPYMPFPATWPQFVPKDKLAGWFEAYVEAMEINFWGTTAFIGGDYDEQSQSWVARVRRG 288
                .***: **  ** *.****:..*:* *...**:  *  *   .. :*: : .*.. : *.

f               -GEDVVLHPKQLVLATGMSGKPNVPSFPGQDVFRGEQHHSSRHPGPDAYVGKRVVVVGAN 347
g               DGSTRTLHPRHVVLATGMSGVPNIPRIAGADTFEGTIEHSSWFVGGREMQGKKALVVGCC 357
e               DGTVRTLRPKHVVIATGVSGIPYVPELPGLSQFAGRTLHSSEYDDANDFAGQRVVIIGTG 348
                 *   .*:*:::*:***:** * :* :.* . * *   *** . .     *::.:::*

f               NSAHDICKALFENGADVTMLQRSSTHIVKSDSLMDLGLGDLYSERAVAAGMTTEKADLTF 407
g               NSGHDIAQELNEQGADVTILQRSSTYVMSSKHGIPGLFGGVYEE----GGPAVQDADLIF 413
e               NSAHDVAQDLHAHGIDVTMVQRSSTTIVSVDPSAAAADASYLTA------PTLEDCDLLS 402
                **.**:.: *  :* ***::***** ::. .       ..           : :..**

f               ASLPYKIMHEFQIPIYQKIAERDRDFYDRLEKAGFKLDFGDDGSGLFMKYLRRGSGYYID 467
g               ASLPYPLLAGIHAGATEAIAEKDAEMLDGLRKAGFKVDFGEDGSGLFMKYLRRGGGYYID 473
e               MATVYPDLYTGSQMITATMKELDKDLVAALNRIGFRTDYGEEDTGQQMKFMRRGGGYYLN 462
                 :  *  :          : * * ::   *.: **: *:*::.:*  **::***.***::

f               VGASELVADGSIHLVSG-QVDHLTEDAVVLTDGTELPADLVVYATGYGSMNGWAADLMGQ 526
g               VGASELIASGEVSVKQGTEIDHFTPDGVVFADGTEMPVDVVVLATGYKNMRESARKFLGD 533
e               VGCSDLLISGQVGLVQYADTAGFVAEGLSLTNGDVVEADAVILATGYQTQQEGVRALLGD 522
                **.*:*: .*.: : .  :   :. :.: :::*  : .* *: **** . .  .  ::*:

f               EVADKVGKCWGLGSDTTKDPGPWEGEQRNMWKPTQQEALWFHGGNLHQSRHYSLYLALQL 586
g               AVADRCQDVWGLDA---------EGELRTVWRRSGHPGFWFMAGNLHQSRHYSKYLAFQI 584
e               EIADAVGPIWGYDD---------EGEVRNTWRRTAQPGLWFSSGNFQLCRIYSKVLAMQI 573
                 :**     ** .          *** *. *: : : .:** .**:: .* **  **:*:

f               KARHEEIPTPVYGLQEVHHLS 607
g               KAQEE-------GLQPIR--- 595
e               RTELDN------G-------- 580
                ::. :       *
```

Legend: N-term. extention

mutated residues

All protein sequences of the cloned *Rhodococcus jostii* RHA1 genes

Rhodococcus jostii RHA1 Type II BVMO:

>1
MSTAAPAVTEVLDVLVVGAGFAGLYQLENLRSRGYSVKVVEAGEGLGGIWHWNRYPGARVDSEGPIYQFTRPDLWDEFAFSELYPGGDELRRYFKYVDAKLD
LSKDIYYNTRIISAEFDDTANTWTVTAENGSSFVCKYFVLCTGFAAKPIFPKLPGMDSFTGINHHTGLWPEGGIEFAGKRIAIIGTGASGVQVAQEASKKAA
QLTVFQRTPVQALPMRQRQLTDEDNAKIKFDLADRFSRRSASFSGFDFDFIPKSALGVSDEERITTYERLWECGFEFWLGTYQDVFVDDDANDTAYEFWRDR
TRARIKDPVIAEKLAPTKKAYPFGVKRPSLEQTYYDIFNQDNVRVVDLHEDPIETITPTGLKTTSEQHEFDIIVYATGFDAVTGGLTAMDIRGTDGTLLRDK
WSNGVRANLGVATAGFPNLLFLYGPLSPSGFCNGPSCAEIQGDLIVNTIDYMRDNGLNRIESEADADAAWSDHVAELTAEALYDKADSWYMGANVPGKPRQL
LNYPGGLPLYLAKWDETVCAGYKGFTLS

>2
MNLPQHVHTLIVGAGFAGMGLAARILQTQPQADVLIIERGDDVGGTWRDNTYPGCACDVPTSLYSYSFAPSADWSHTFARQPEIHRYLKKVAADTGVRSRVV
TDCELQEAHWDADEAVWTVRTSRGTVTADVVVAATGALSTPSVPDMPGLETFGGTTFHSATWNHDHDLTGERVAVIGTGASAVQFVPEIAPVAEHLTVFQRT
PAWVIPRLDRELSGSEKRLYRRLPLVQKAVRGTVYGFREALGGVLAHATGLLPVFEMVAKAHLRRQVRDPELRRKLTPNFTIGCKRMLLSNDWLRTLDRPDV
TLVDAGLAGVTPDGVVDALGNEHKVDTIIFATGFTPTEPPVAHALRGANGRTLAEHWDGSPSAYKGTTVAGFPNLFLMYGPNTNLGHSSIVYMLESQSAYIA
DALNVMHHSEITAFEVTEEAQRRYNTRIQSELQTTVWNKGGCSSWYYDSEGRNSVQWPTFTWKFRSQLQRFDQENYVSRRRAAKETVA

>3
MNTTLLEELDVLVIGGGFSGVYQLDRLRTLGYNVKIYEAGTGLGGVWHWNSYPGARVDTWAPVYQFSREELWRDWNWSEMYPGRDELVRYFEYVDEKLDLSK
DVRYETRVLAGRFDEETHRWTLVSRNERTGEEFTTQAQFVIMCLGAGSKPLFPNIPGLEKFGGDCFHTARWPLEGYDLAGKRVAVIGTGASGVQVIQEASKV
ADHLTVFQRTPNTALPMNQRALGEADNAEMKKTYPERFANRKNTWAGFDYDFLKENIQDLTEERRNEILEELWTNGGLQPWLGGFLNVLFDKDDNDILYAFW
RDKTRQRITRPELVELLAPTEPIHPWGVKRVSLEQNYFESLCRDNVELVDTSANPIREVASDAIITADGTRHEVDVIVLATGFDSVTGGLTAIDIRGTGNET
FEEVFRGGSRTALGKATVGFPNLLYVYGPQSPNAFCNGPTCAELEGEHLIQIVEHMRNNGYTRIEAKPEAQQYWGAHIAELTSATLFPLAKSWYMGANVPGK
TVEMLMYPGGLSVYLEILEKAAAGGYQEQFELV

>4
MTAIHAPKNDLAQSAETPTYDVIVVGGGFGGIYQLRHLRDRGFSVILLEASGGFGGAWSLNRYPGARVDSHAPVYQFTDEYLWKDWDFSQMYPDHEEMRSYF
NYVDSKLDLSKDSRFNTKVVGATFDEEQRMWSLETQDGATFRARFVVFATGSTTEPYTPSIPDMDAYQGELVHTARWRSDLDMTGKRVAIIGTGASAVQVVQ
EAGPVVENLTVFQRTPNISLPMQQKYLDDEEQAALKNKMPDVAAKCRETHAAIDYDFDPRSGFETPEDERNAVFERLWNQGGFAFWLGNFSDYLFNDKTNAL
TYEFWKNKIKPQIKDPVKAELLVPEIAPHPFGAKRPALHQNYYEVMNQTNVSLVSTKETPIVGFTETGIRTADGVEHGEFDIIVLATGFNNNTGALTSIDVQ
NANGVTLRDKWSQGVDAYLGAVTAGFPNAIFVYGPQSPAAFANGSTNAELQGEVMVDFFEFLRSNGLTRFESTVEADKAWTAHINETDDTALFNRAKSWYNG
GNIPGKKMQMLQYLNGVPTYLQFWQKEKESGYTDGLTVS

>5
MGGEVHRCRVLVIGTGFSGLGTAIQLRRRGRDDFILLEKAREVGGTWRENTYPGCACDVPSHLYSFSFEPNPDWTRMWSGQEEIFDYLRGLADKYDLRRNIH
FGRTMTGGHWDAERRRWHVHTESGDEYVAQFVVSGIGALHIPNVPDLPGADTFDGTEFHSARWNHDYDLRGKKVAVIGTGASAVQFVPEIVGDVAELHLYQR
TPPWVIPGLNFGIPPEARRLFGRVPLARRMVRAAVYWTYESLALGFNGHSRLMRPIESMARKNLNRTVADPELRRKLTPSYRIGCKRILGSDVYYPALISPK
TVVITEGIAEVRPHSIVAGDGAERVVDAIIYATGFHVTDGFDNVELTGVGGRRLADEWEEHGIRTHLGITVAGYPNAFFLFGPNTGLGHNSVVFMIESQIRY
ALELMDLVDRRGADSAAVRPAVQSGFNADIQRKLARGVWSTGGCVSWYLDSHGVNRTIWPGSTVRYWQRTRSVEPADFEFTPG

>6
MTHYDILIVGAGISGIGAAIRLKQSGIDNFAILEKGDALGGTWRDNTYPGCACDVPSALYSYSFAPNREWSRLFAGQDEIRRYIERTAAEHGVPAHVKFGTE
MQRAQWSEQSRRWTVDTSAGTFTANAVIAAAGPWNEPLVPTVPGLDTFTGEVFHSSRWNHTYDLTGKRVAVVGTGASAVQFVPAIQPTVESLHLYQRTAQWV
LPKPDHTLPGVERAILRRVPGAIRALRRVEYAIMESLGLGFRHPWILRVIQQVGKAQLRAQVRDPKLRKALTPDYTLGCKRLLLSNTYYPALTRPNVEVHAN
AVESVRGNVVVGSDGAEREVDAIIFGTGFHILDMPIGSKVFDGDGRSLDDHWKGSPQAYLGTTVAGFPNAFVLLGPALGTGHTSAFMILEAQLDYLIQAVTA
ARSNGWTRMEPRREVQDAFNAQVQEALATTVYNAGGCQSYFLDVNGRNSFNWPWSTDRMRQRLGRFDEAAYDVSREPASATATSR

>7
MVSYMSLPVTDTSAPPAGVRHIDTLIIGSGFAGLGAAIKLTQAGKTDFLVLERGSDVGGTWRDNTYPGAACDVPSHLYSYSFALNPEWTRSFSTQPEIQKYI
QSVADKYKVRNKHLFGCDVQSAHWNESTTRWEVTTTKGNFVAKVLVSAVGALCEPSLPDIKGIEGFEGEIFHSARWNHDADLTGKRVAVIGTGASAIQIVPA
IGKKVSHLDVYQRTAPWILPRADREYTKLEHTAFKYLPGFQKLCRTGIYWMRESQVVGLAKAPVFMKPLQFAAERHLRRQIKDKALRKKVTPNFQIGCKRML
ISNNYYPTLAQDNVDLVTDGIAEVTADSVVSKDGTVREVDAIVVATGFHVTDSPTFEGIFGKDGRSLADVFDEGGQQGYKGAAIANFPNMFFLVGPNTGLGH
TSMVFMIESQLNYLVDALQTLDKYDIGKIEVRQDAQDRYNAELQEKLSHSVWNNGGCASWYLDKHGNNTTLWPGFTFQFRNETKRFDLTAYDSVATADLPAP
VHVNGKTPGPAAIPAQIDLDDDKVTAQ

>8 cured
MITEIRKRAREVLTNDLTAVLGVPDPELFRRMASLRTAGTVAEEFVPILLEQAGFQIGQRRVPVTKKPPADLGVIVIGAGMIGLNAAIKLGEAGFGYRVFES
RDDIGGTWSRNVYPGAAVDTPSHYYSYSFELNPDWSRYYPTGPEYLDYMHAVAEKYDLYKNIELSTAVVTAEWDEAAQKWTVVTRRTDGSTPRHRASAAITG
FGFLNSPNKTDIPGLDTFEGTVVHTAEWDPSLDLTGKRVVVVGTGCTAVQIVASIVDDVATVDAIVRSPHWIVPEKLWSTRYRPGRSGQCGTCRTSRIGSGS
EPTGLPQNNLYMMPRIDPEWAATHLSVSPVNDLVMQTSLQYLEQTLPDRPDLREKLTPSIRPYAKRIVKDPGFLEALERDHVSLHRASPKHVHPGGVSLSSG
EFVEADVIVLATGFKVEYASFIDITGRNGKKLADKWDHGQDPRAYLGIQVSGFPNLFVTAGPNAAPNHGAGHNITSEERVHYIVECLQYLVENDFSAMDVKP
EALTVYNEKVDEALDQTVWAHPGEGVTGYYRNQQG

>9
MTVQDNDFDAVVVGAGISGLYAVYKLRQRGMRVHGFESAEGVGGTWYHNRYPGARCDVESIDYSYSFDEELQQEWTWTERFATQDEILRYLEHVADRHDLRS
AYDFLTRVTSATYDEETTRWSITTDTGQNVTARFCVLATGVLSATNKPDIPGRDTFGGATYHTGEWPHEPVDFAGKRVGVIGTGSSGIQSIPVIAEEAAEVF
VFQRSPNYSIPAGNRPLTGEYIAEVKANYAERRRLSRMSGGGTPNSAYPKGALEVDAEERRRVYDEWWQRGGYLFAKAFPDQTISQAANDTAREYVEAKIRE
MVTDPDIADQLVPTDHPIGTKRIVTDNGYFKTFNRGNVTLVNLRRTPITEITEAGVLTTNSFYGLDMLVFATGFDAMTGSLSRIDIRGRAGRNLRDEWSAGP
RTYLGLSVAGFPNMFILAGAGSPSVLANMVLMAEQHVDWISNCLDYLDEHNIETIEATDESVDEWVAECNEKAAGTLFPTADSWYMGANIPGKPRVFMPYIG
GFGNYNTICAEVAAAGYKGFELGERRVDVDQ

>10

MTETIAAGLAVPSDRDAQLYNAIAESDPAPLLMALVHATGDTGLLDEFGARLTIEEPGNHYRTGIRPTAPPGIYPEDVAEDIRIRAREILTPDVVAELGVPD
DELFVRMATVCTSQRVDAEFAPILLEQAGFTKNRRHVPVTVAPPADFDVIVIGAGIVGINAGIKLGEAGFRYTIIEEREDVGGTWYRNTYPGAAVDTPSHYY
SYSFELNPNWSKYYPTGPEYQNYLLDVVEKYRLREHIRFRTRVLSARWLDDEHRWEVVTEDGEGSVVRHRARAVITAMGMLNAANIPEVDGLDSFAGRVVHT
AEWDSDLDLSGKRVVVLGTGCTSVQVVANIVDQVEALDVVVRSPHWLVPEKAVSGDVTEGEKWALANLPFYDRWFRLRSYWFASDNLYPLPRIDKEWAATHL
SASPANDMVLRTAQEYLQTSFTDRPDLIAKLTPDFRPYAKRIVKDPGFFAALNREHVSLHRASFEKVTPEGVYTTEGAFIPADVIILATGFKLQFTTSIEIE
GRDGRTLSEVWNGGDDPRAYLGVQVAGFPNLFITAGPNSAPNHGAGHNILSEEHVHYIVECLQYLLENGHDAMDVRQDVLDTYNRKVDAALDDTVWVHPGAE
VNGYYRNSSGRAIVPCPWRLVDYWTMLREPNPDDLTFIGRRAEGRREASAR

>11

MTTASIDTRELDEANGVLDVLVVGGGFAGLYQLDQLRSRGFSVKVVEAGDSLGGIWYWNCYPGARTDSTGQIYQYSREDLWKDWSYDELYPSWSGVRDYFAY
VDRKLDLSRDIIFSTRVTSADFDGERNQWTVRTDTGRMLRARSVVICTGFGAKPHIPSINGLNSFAGESHHTALWPQEGLDMAGKRVGIIGTGSSGVQVTQE
AAADAEQITIFQRTPNLALPMRQQQLTGQLKEKLKENLPERFAQRRRSFAGFDMDFIPKSVFEVSDEERADTYERMWATGGFELWLANYQDILLDERANRIM
YDFWRDKVRQRVTDPVKAEKLAPMDPPHPFGTKRPSLEQNFYDVVNQENVDIVDVNEDPIERITPAGVQTKSGLHEFDILVFATGFDANRGGITSIDIRGTN
DQLLSHKWSERLDTFMGLTTAGFPNLMFVYGPQSPAGFCNGPTCAEVQGEIVVDFLTHVRDGGYQRFETSEDAEQSWTAHVEEVFHMSLFPRAKSWYHGANI
PGKPSQMLNYSGGLPSYFDHWEENVAAGYKAFTLS

>12

MVRSAIPVELPVDSVDHPPEVVDVLVVGAGFGGLGTAIRLKQAGIDDFVVLDRAEDIGGTWRVNTYPGAQCDIPSILYSFSFAPNPNWTRLYPLQQEIHDYL
RSCAENFGIVPHLRMGHDVQDAAWDDDSQVWHVTTSRGTWEARILVGAMGPFSEPAVPNLPALESFRGAVFHSAAWDHEHDLAGERVAVIGTGASAVQIIPR
IQPIVGSMTVFQRTPTWILPHPDQPMTGWPRKLFARVPVAQRLARSGLDLVQEAMVPGFVYKPALLKGLAALGRAHLRRQVHDPELRTKLTPTYAFGCKRPT
FSNSFYPALAQPNVDVITDGIREVRSNGIVTEDGVLHEVDTIVMGTGFRLTDNPAFDVVRGRDGRTLAEAWNGNARAYLGTTISGFPNFFMLLGPNSVVYTS
QVVTIEAQIAYILSCLQEMNAQGAASIDVRPEIQQAFVDEVDERLQTSVWNTGGCNSYYLSGEGRNFTFYPGFNRRFRARTRRVDLHHYIISGAGASAKSIV
RTAG

>13

MKHITTSHDSDQETPVTITTIPEVETMDFDVLIIGAGISGIGAAYHLKTRRPDTTFAILEGKDAIGGTWTQFRYPGIRSDSDMPTFGFGFKPWTHKKAIADA
HIILDYLQETVTENHIDEHIRFGYRVSSAEFSSSAGRWTVTAQRSGSDETVQITARFLFSGTGYYNHEAGFTPEFDGIEDFTGQVVHPQHWPEELDYSGKKV
VVIGSGATAATLIPAMAGTAAHITMLQRSPSYVLSLPAEDAIANTLNKLIGPKRAYPIIRRKNIMMHRGIFKACRRSPKLMRKLLIANARRQLPKNFDVDTH
FTPRYNPWDQRLCMVPNGDLFKAISTGHASVVTDRIERFTATGIRLESGQELEADIVVTATGLNMLAFGAIQLSVDGEPVNPPDTTVYKSMMLSGLPNFVFA
LGYTNISWTLKVDLISEHFCRLLDHMDERGYTTVEPVLTDPGMERVPLLDLTSGYVQRAVAAFPRAGTSGPWTAAMAYEKDVERLREGPIEDADLRFTANQP
ALLAS

>14

MSKTISADVDVVVVGAGFAGLYALRKLRDTMKLSTRVFEAGSEVGGTWFWNRYPGARCDIESVHYSYSFDEDLQQEWQWSERFAGQPEILRYLEHVADRFDL
RKDITFDTRVVGVHWDDENSVWTVRTDDGAVVRSRYFISGAGNLSVPKTPEFGGIDNFRGEVLLTGNWPREGADFTGKRVAVIGTGASGIQAIPFIAEDAAE
LVVFQRTPNFATPLGNGPMDPNELADIKSNYADVRTAARNHFLGVPFNQVQPSALAVDAEERRRTFDERWNAGGFRLFIDSYQDILFDKKANDTIADYIRDR
IHERVQDPAKAATLAPTGYAYGTKRPPLETNYYEAFNRDSVSVVDVKSTPIDEITPTGVRVGDRVYEVDTIVLATGFDAMTGPLMAMDIRGRGGLPLAEKWE
HGPRTYLGIMVNEFPNLFLITGPQSPSVLYNMPLAIEDHVDFATDAIDYLDRRDLDVIEPTAQAESDWGALTNEIADQTLLPETNSWYMGANIPGKPRACMV
YLGGAPTYRATCDEVVAGGYSGFALTRAEARAASTVS

>15

MSARTEVDARANRIGDVDAVVVGAGFAGLYAVHKLRSLGLTVQGVEAAGGVGGTWFWNRYPGARCDVESVDYSYSFSRELEQEWDWSEKYATQPEILAYINH
VADRFDLRDRFLFGTRVTSAELDEESLRWEVRTDRGDVLSARYCIFATGALSTANMPNIAGRESFTGDTHHTGQWPHEGVDFTGRRVGIIGTGSSGIQSIPL
IAEQAEHLYVFQRSANYSVPAGNQAWDDEMRRAIKAGYEERRRLSRESGGGSPYNAHPKSALDVSDEERREAYETRWKLGGVLFAKTFPDQTKTEAANATAR
EFAEEKIRLLVDDPAVADKLIPNDHPIGTKRIVTDTHYFETYNRPNVTLVDLKAAPIESITPSGITTADADYALDTLVFATGFDAMTGALDRMRIVGRGGVP
LSEYWSEGPKTYLGLGVPGFPNLFVVTGPGSPSVLANMVLGAEQHVDWIADCIEHLWEKDYDAIEASVPATEQWVEHCRDLAAQTLFPLANSWYMGANIPGK
PRVFMPYLGGFGAYGRICADVAEEGFRGFEFSRSRTRLADPVG

>16

MSHTETAAETTGAKTPVEHVDVLIIGAGLSGIGAAYHLQDNFPRRTYAILESRESIGGTWDLFRYPGIRSDSDMYTLGYRFKPWSGEKSIADGPSILEYVKD
TAAEHGIDRNIRFRHKVVRAEWSTADSHWTVDAERTDTGETVRLTADFLMSCSGYYRYDEGYTPEFPGLDRFGGRVVHPQQWPEDLDYEGKRVVIIGSGATA
VTLAPSMAADAAHVTMLQRSPTYIISMPAKDKLANKLRRHLPAKLAYGLTRLKNASVATAIYQLCQRYPEFMKGRIRQLQEKWLPKGYDIDTHFTPRYNPWD
QRLCLVPNGDLFRAIRNDEVSIVTDHIDTFTETGITLKSGEELHADVVVTATGLNLLAFGGMTLAVDGHDIDLTETMAYKGMMLSGVPNFAFVIGYTNASWT
LKADLVCEYVCRLLAHMDANGFTQCAPERDSSVEEEPFLDFAAGYVLRSVESFPKQGSKAPWRLRMNYFRDLVALRHGKILDDAMTFSRP

>17

MTTGTTEPDVLVVGTGFAGLCMAIKLKEAGEENVVVLEKADRVGGTWRENTYPGCGCDVMSLMYSFSFAPNRKWTRMYARQPEILDYIERVVRDYDLAPHIR
FGAEVISYEFDETTDRWRVETRSGSVYHPRIVVAGPGPLHKPSVPDLPGRKSFSGVAFHSAEWDHSVDLTGKRVAVVGTGASAVQFVPEVAKTAAHVDVFQR
TPHWILPKLDRPITAGEKAVFKAVPGVQKAYRGAIYWSHESLIAGFLHPRLMTVLESAARGLLRRQVRDPELRATLTPDYIIGCKRILVSSNFYPALQRGNV
DLVTSGISEVTERGIRTDDGTMHEADVIVYGTGFAAGDRFENEHIVGRRGLTIQRAWRDGMEAYLGVAVAGFPNFFLMMGPNSGGGNQSIVFVIEAQAHYIT
RCLALMKKRDATRIEVRAGAQREFNRVVHRKLAGSVWNSGGCDSWYLDSTGHNRAAWPGSSASYWRRMRTPDDRHFELSSLAEREDDTEYRGPGVLTSGDLT
VAVEVFLNGHIEPLDGLYHWYGRVVGDGVDAAKGRNRTPLFLTIGDGPEVPAALAERDPWGHFRIAGVGTPPFPLAPVEVEVPISRAKLASAE

>18

MKAAQQVQEAGAATDDRINPDRLTESEIRTAVARANVPSLLMVVFQTTGDEKWLAAPYRPTRGKGLGDHDSGGLEEPIQDEIREAAVKAILDLQNGALPAVE
TPSPELTVRMISVCTGEEVGEEYGPMLSLELARRAAPDAPSLALEPVDAPEGYSVVVIGTGVAGIAAAQQLEDMGIDYVILEKQPEAGGNWWQNTYPGAGVD
TPSHLYSFSFAKNDWTTHFELRNELQAYFGAVLKDLGAGERVRYGTEVRSTRYDEAAAQWSVDMINPDGSSSTLRADVVISAVGVLNRPKTPNVPGMDSFTG
TSFHSAAWPDDLDLDGKRVAIVGTGASSMQIAPAIADRVAHLSIYQRSPQWVAPFEKFRAPIPMELRRLMQTCPIYHSWYWIRLFWQFGDKVIESLRVDPEW
EHPERSVNARNDAHREYFTRYITSQVGDRTDLLDKVMPDYPPFGKRILLDNGWYSTLRKDNVDLVDRSVTAVRPEGLVDDQGAENDVDVIVWATGFEAARFV
SSMDVVGMDGRTLREVWNDDDPKAYLGVSVPGFPNFFMLGGPNSFPGSGSFMFFMEVQMRYIRGLLTEMFKKGIKAIDARPEANEEYNELVDSTHARTVWTH
RGMSTYYRNSHGRVVFVMPFLNVEYWQMTRRPDLENYTAR

>19

MTATTTQHAAAPDGGGDERHLRVVVVGAGLSGIAAAVKLERAGITDFVVLEKSDRVGGVWRENTYPGCGVDIPAPVYSFSFHPNPRWRSNFALQPELLSYIE
DTVDTFGLQSRISMQTDVREAAWSDERRRWILDTSRGTIVAQHVIFAAGPITEPSTPAVPGIDRFDGDVFHSARWNHDVDLTGKRVAVVGTGASAVQFIPEI
QPDVEELYVFQRTPAWVVPRLDFPFPRIAQWAFARVPAVQRALRRLLDVILRTLTWVMRCERTARLLNPIGTRWLARQVPDPALRAALTPNFTLGCKRLLLS
NTYLPALTKSNVELIPHALAEVDGRVVVGADGTRREVDVIIFGTGFDVSHPPIASRIRGRDGTLLSEKWSKSPEAYLATTTPGAPNAYIMLGPNILVYNSFL
GLAETQLDYVIDGLTTAERQGIEVLEVRDQPFRRFNDAVQKGLEPTVFNNGGCSSYYLDADGRNFAAWPWSTGSLRRRLARFDLENYAIRPYRTEQSPALHP
SGKSR

>20

MTASQADTATRTGKHSNNDVLDVLIIGGGFSGLYALDRIRDLGFTAKVWDAAGGLGGIWWWNCYPGARTDSTGQIYQFSHKDLWKKYDFAELYPGHDGVRNY
FEYVDSQLDLTRDVVFDTFAESCTWDEETRQWTARSADGKVQNARQVIVATGFGAKPLYPNLEGLDLFAGDCYHTARWPQEGVDMTGRKVVVMGTGSSGVQV
VQEAGHVAEHVTVFQRTPNLAIPMQQRALTHDDNEQFRKGLPERFEARYKAFAGFDFDFLPQNAADLSMEERDAIYEKMWAEGGFEMWLGNFQDILVDEDAN
RTFYDFWRNKVLERVTDPKKAAIVAPETPPHPYGVKRPSLEQDYFDVINQSNVEVIDSNLTPIRRVLPHGIETDDGVIECDLLVLATGFDNNSGGIMAIDIT
GVDGLSIQDKWKSGVDTCMGLSTRGFPNMMFLYGPQSPSGFCNGPTSAEYQGEIVVEFLQHLRDNGITRFENTEESEKQWRAHVDELFVNSMFTKARSWYWG
ANVPGKPAQMLNYSGGVPQYFARWDKIKANGYAAFETN

>21

MSTTASAPTETSATESLELDALIIGAGVAGLYQLHQLREQGLRVRAYDTAGDVGGTWYWNRYPGARFDSEAYIYQYLFSEELYKNWSWSQRFPGQPEIERWM
HYVADTLDLRRDIQLSTMITSAHYDERADKWIVRTDRGETITTRFLVTCSGMLSAPMSYVFEGQEEFSGPIFHTSRWPKEGADLDGKRVAVIGVGATGIQVI
QTVADKVEHLKVFIRTPQYALPMKNPTFDESDVAAYKSRFAELKETLPNTFSGFEYDFEHVWADLTPEQRNDVLEEIYENGSLKLWLASFGEMFYDEEISEE
ISEFVRRKMRARLQDPHLCDLLIPTDYGFGTHRVPLETNYLETYHRPNVEAIGVRDNPITRIVPQGLVLADGTLHEVDVIVMATGFDAGTGSLTRIDIRGRG
GRALKDDWNRDIRTTMGLMVHGYPNMLTTGAPLAPSAALCNMTTCLQQQTEWIAECIRYMRAHDHTVIEPTLAGEDEWVAHHDETANATLVSKTDSWYNGAN
VPGKPRRVLSYIGGVGTYREKTLAAAAAGYKGFQLS

>23

MSTTASAPTETSATESLELDALIIGAGVAGLYQLHQLREQGLRVRAYDTAGDVGGTWYWNRYPGARFDSEAYIYQYLFSEELYKNWSWSQRFPGQPEIERWM
HYVADTLDLRRDIQLSTMITSAHYDERADKWIVRTDRGETITTRFLVTCSGMLSAPMSYVFEGQEEFSGPIFHTSRWPKEGADLDGKRVAVIGVGATGIQVI
QTVADKVEHLKVFIRTPQYALPMKNPTFDESDVAAYKSRFAELKETLPNTFSGFEYDFEHVWADLTPEQRNDVLEEIYENGSLKLWLASFGEMFYDEEISEE
ISEFVRRKMRARLQDPHLCDLLIPTDYGFGTHRVPLETNYLETYHRPNVEAIGVRDNPITRIVPQGLVLADGTLHEVDVIVMATGFDAGTGSLTRIDIRGRG
GRALKDDWNRDIRTTMGLMVHGYPNMLTTGAPLAPSAALCNMTTCLQQQTEWIAECIRYMRAHDHTVIEPTRAGEDEWVAHHDETANATLVSKTDSWYNGAN
VPGKPRRVLSYIGGVGTYREKTLAAAAAGYKGFQLS

>24

MTTSMKAANPMNFPSTSDTGIVDVLGVGAGFSGLYLSHRLTTAGWTFAGFEAGPSVGGTWFWNTYPGARCDVESIYYSYSFDEALQQEWTWSQRFAPQAEIL
SYINHVADRFDLRKHFTFNTRVVGATWNAAERLWEVQLDNGETRRGRYLISGAGGLSTPKDFDVPGLGNFTGLQVSTSRWNISLDDLAGKRVAVIGTGSSGV
QAIPLIAEVAEHVTVFQRTPNYVMPARNAELPLERVDSIKDDYPAIREECRHSPGGIPDRPVTDKAFDVSAEERQRRYEAAYERSGFNGVGGEFADLLTDVE
ANRTASEFIHDKIREIVEDPATAELLVPRYHPLGAKRSVFGTDYYETYNRPNVSLVSLRDEPIETMTANAIVTSKGTYEADAVVLAIGFDAFTGPLYGLGLT
GASGRKLQETWQDGIRTYLGMMTTDFPNFFMVAGPQSPALASNVVMTIEQAVDWIADLIEHARDSGATLVEATPEGQNDWVDITEETVAQTLYATTDSWYRG
SNVEGKPNTFMGYVGGVGKYRRMCTEIAKRGYPGVRIDGETESPHLGPIHREIS

Rhodococcus jostii RHA1 Type II FMO

>a

MSEHQVAIVGAGTSGVAAAVALADRGINPLLIDRADQVGSSWHSRYDRLRLNTGRQFSHLPNRPYPKGTPTFPTREQVIEHLERHARADGIELRLGCPVERL
DLTDGHWRLTTAAGSVDAAEVVVATGFDHEPFVPDWPGRGDWRGALVHSSQYRNPSQYNGKRVLVVGAGCSGMEIAYDLATGGAAKVWLSARTPPNIMLRQG
PGGIPGDFIATPLYHAPVPIADAIARFGRERSIGDLREFGLPIPDEGIFARSARLGVAPAIVDKELIAAIRDRSIEVVRGVESLDADSVWLVDGVRIDPEAM
VCATGFRQELEKLVGHLGVLDERGWPHATGEKPAAERLRFIGFVPRPSQIGFAAKQARRAARAIARELR

>b

MTEQHTVVVIGGGQAGLSISWHLVQRGIDHVVLERESIAHEWRDSRWDSFTLVTPNWQCTLPGYTYSGGDPDGFMNREQTYQFVRGYADTFDPPVREGVAVV
AVRQSGSGGFDVDTTEGPMHADHVVVAVGGYHTPVVPRFAERLPADITQLHSSQYRSAGALPAGEVLVVGNGQSGAQIAEDLHLAGRTVHLVTGGAPRVARF
YRGRDCVAWLHDMGTYDVSIADHPGGLGKRENTNHYVTGRDGGRDIDLRAFALAGMRLYGRLLDVVDGTLRFAPTLESSLDAADAVSESIKDSIDAYIDRAG
IDAPREERYVPVWRPEREVTELELPTSGITSVVWSIGFRTDYRWLHAGVFDGEGHPTHNRGVTAVPGLYFLGLPWQHTWGSGRFAGVARDAEYLADRIELEA
GVLPATATLA

>c

MSTERFETIVIGAGQAGLATGYHLTRCGQRFVILDAHDRVGDVWRERFDSLRLYSPARYDGLPGWGIPAPAWSWPGKDEVADYFEAYAQRFALPIRTGTTVD
GLSRDGDRYVVTAGTDRFEADNVVVASGTWQSPVVPDLAERLDPRIRQLHSSDYRNPSQLQDGPVLVVGCSHSGADIALEASRSHRTTICGPVRGEVPFDIE
GRLAHLAVPIMWFMANHVLTERTPVGRKMCTHVRSGGGPLLRVKRADLAAAGVEHFPAKVTGVHDGRPVLDDGTAFDVRNVIWCTGFRKDTSWIQIPVTGSD
GWPEQSRGVSPDHPGLYFVGLPFLQAFASMLTGGVGRDAAYVAKHIAKRVVVRSPEAVA

>d

MNSEVDVAVIGAGQAGLSAAYYLRRFGVEPESGFVVLDHAPGPGGAWQFRWPSLTLSTVNGVHDLPGLGFADTIGVDPNDPEAALVHAASAVPQYFATYEKQ
FELPVHRPVHTRVVCARDERLRIETDRGVVSARGLINATGTWERPFIPRYPGAESFTGRQVHTKDYSSAQDFAGQHVLVVGGGISAVQLLDEISRVTTTTWV
TRRPPEFRDEPFTPEIGRAAVALVEDRVRRGLPPGSVVSVTGLPVTPAIRAARERGVLARQPMFGEITADGVRWPDGRELKVDVILWCTGFRSSLDHLAPLR
LRGPGGGITMTGRLATQVASDPRIHLVGYGPSSSTIGANRAGQAAARELTRHLGLSAGSARPT

>e

MTTTFSDTDLRTDAQAWLDGFSRFLAAELAPTAVFAPQAYWRDVLAFTGDLRTFSDEIPAELLRRQELTKATNIRIAEDRTPPRLVERAGIPCLEVIFEFDT
LAGSAVGVARLVDVPERGLLVRSLFTTLDQLADHPERTGEHRPVGQADSSKFGGPNWLDRRIAAQAYENRDPDVLIVGGGQSGLTLAARLGQLDVDALVVDT
HARPGDNWRTRYHALTLHNAVWLNDLPYMPFPATWPQFVPKDKLAGWFEAYVEAMEINFWGTTAFIGGDYDEQSQSWVARVRRGDGTVRTLRPKHVVIATGV
SGIPYVPELPGLSQFAGRTLHSSEYDDANDFAGQRVVIIGTGNSAHDVAQDLHAHGIDVTMVQRSSTTIVSVDPSAAAADASYLTAPTLEDCDLLSMATVYP
DLYTGSQMITATMKELDKDLVAALNRIGFRTDYGEEDTGQQMKFMRRGGGYYLNVGCSDLLISGQVGLVQYADTAGFVAEGLSLTNGDVVEADAVILATGYQ
TQQEGVRALLGDEIADAVGPIWGYDDEGEVRNTWRRTAQPGLWFSSGNFQLCRIYSKVLAMQIRTELDNG

>f

MTQTVQPAAVQTSLTPQERVDLWLASFESALAARDVDRAAGMFAVDSFWRDLVAFTWNLKTVEGRDAVAAMLHARLDDTDPVNFRTTETPDEADGVTSAWIE
FETATGRGKGHLRLKGDEAWTFLTTMQELKGHEERRGRNRVKGAVHGSGGDTLSWAEKREIEERELGYTRQPYVLVIGGGQGGIALGARLRQLGVPAIVVDK
NERPGDQWRNRYKSLCLHDPVWYDHLPYMPFPDNWPVFAPKDKIGDWLEMYTKVMEIPYWSSTTCTSATFDDETKEWTVVLDRDGEDVVLHPKQLVLATGMS
GKPNVPSFPGQDVFRGEQHHSSRHPGPDAYVGKRVVVVGANNSAHDICKALFENGADVTMLQRSSTHIVKSDSLMDLGLGDLYSERAVAAGMTTEKADLTFA
SLPYKIMHEFQIPIYQKIAERDRDFYDRLEKAGFKLDFGDDGSGLFMKYLRRGSGYYIDVGASELVADGSIHLVSGQVDHLTEDAVVLTDGTELPADLVVYA
TGYGSMNGWAADLMGQEVADKVGKCWGLGSDTTKDPGPWEGEQRNMWKPTQQEALWFHGGNLHQSRHYSLYLALQLKARHEEIPTPVYGLQEVHHLS

>g

MTATLDAPVDTNIPQPGDIARRWLAGFGATLERGDARGAAQHFLVDGWWRDLLSFTWDLHTTHGRADIESRLADSVPVHEPRHLVLSPAHPAEAVADPEGDW
IQAFFTFETTLARSRGFVRLRRDDDGEWRAWTLISAMEEIKGHEEKKGHRRVQGTNHGAHRGKINWLDRRTAKGEFETEQPAVVIVGAGQGGLALAARLGQL
GVDTLLVERNDRIGDSWRKRYHSLVLHDPVWYDHLPYLNFPDHWPVFTPKDKLANWFEFYADAMELNVWTGTEFTGGSYDDATGEWTVTVARDDGSTRTLHP
RHVVLATGMSGVPNIPRIAGADTFEGTIEHSSWFVGGREMQGKKALVVGCCNSGHDIAQELNEQGADVTILQRSSTYVMSSKHGIPGLFGGVYEEGGPAVQD
ADLIFASLPYPLLAGIHAGATEAIAEKDAEMLDGLRKAGFKVDFGEDGSGLFMKYLRRGGGYYIDVGASELIASGEVSVKQGTEIDHFTPDGVVFADGTEMP
VDVVVLATGYKNMRESARKFLGDAVADRCQDVWGLDAEGELRTVWRRSGHPGFWFMAGNLHQSRHYSKYLAFQIKAQEEGLQPIR

Other cloned *Rhodococcus jostii* RHA1 flavoproteins

>NMO H, accession number RHA1_ro08654

MLEHLDLVGIGAGPSNLSVAALSAPVGRLRCKFLDRQPTQRWYPGLMLSAAVLQVSHLKDLVTLVDPTSRYTFLNFLARTGRLHRFASLHTPLIARREYESY
LRWVSDQLDEVQFGCAVEEVTFDGQAFRVESTRGTYAAQHLSIGVGPRPYVPELATGTLGEDVFHSSDFGYHTDSLAGRDVVVVGGGQSGAEVVEHLLQRSG
RDAVGSLTWASRRIGFQPLDESPFTNEWFHPDYVRYFHGLSQSRRSQLLDAQQLASDGISKGLLESIYRRLYYNDFVDSDRIRTTLLPGRELTGLCRGPGGK
GWRTTLTHIDTGEIDSVGADIVVLATGYHFPLPEFLHTLGGRIARTNCGLPQLAADYSVSWAGPAGNKMFFLNAGKLSHGIADPNLSLASWRAATVLNTITE
TPLYPDLRSSTCSWDVADRAATHPPVDSGVDLLTESSRQ

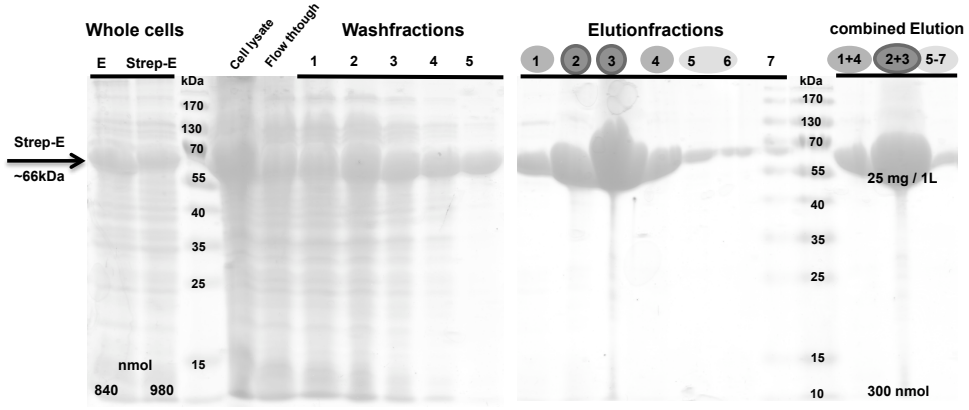
> putative tryptophan monooxygenase accession number RHA1_ro03531

MTIPVTPDSTAADTDPRPLTMFGPDFPFAYDDYVTHRSGLGAVPTDRHGTEVAVIGGGLSGMVTAYELMKLGLKPIVYEADQIGGRMRSHPFDGHPDVVAEM
GAMRFPPSSTTLFHYLDAMGLRTEAFPNPLADSTPSTVIDLKGESYYAQHLDDLPPVFREVSDAWQRTLEDHAELIPLQQAIRDRDTAELKRIWNDLVVRLD
DQTFYGFLASSKHFASFRHREVFGQVGFGTGGWDTDFPNSILEILRVVITAADDHHRGIVGGSQQLPLRLWTDTPTGMAHWPDGTSVRALNGGATRPRVTEV
RRTAPHQVTVTDSSGEIRTYPAAVFTAQSWMLLNNIHCDDDLFPIDHWTAIERTHYMGSTKVFALVDRPFWKDKDPVTGRDLVSMTLTDRMSRGTYLLDHGD
DKPGLICLSYTWSDDSLKLLPLDATERMNLMIKSLEEIYPGVDFRSHIISTPVTLSWETERDFMGAFKANLPGHYRYQERLFSHFVQDELHPRHRGIFLAGD
DISWTAGWAEGAVQTALNAVWGVMHHLGGAAPADNPGPGDRYAELGPIRLCD

>putative putrescine oxidase, accession number RHA1_ro05606

MPTLQRDVAIVGAGPSGLAAATALRKAGLSVAVLEARDRVGGRTWTDTIDGAMLEIGGQWVSPDQTVLISLLDELGLETFDRYREGESVYISASGERTRYTG
ESFPVDETTRKEMDRLIQILDDLAAQVGAEEPWAHPLARELDTISFKHWLIEQSDDAEARDNIGLFIAGGMLTKPAHSFSALQAVLMAASAGSFSHLVDEDF
ILDKRVIGGMQQVSIRMAAALGDDVFLNAPVRTVQWSENGAVVLADGDIRVEASRVVLAVPPNLYSRISYDPPLPRRQHQMHQHQSLGLVIKVHAVYETPFW
REDGLAGTGFGASEVVQEVYDNTNHEDTRGTLVAFVSDEKADAMFELSEEERRATILGSLARYLGPKAAEPVVYYESDWGSEEWTRGAYAASFDLGGLHRYG
KDTRTPVGPFHFSCSDIAAEGYQHVDGAVRMGQRTAADIVARLGK

Purification of Type II FMO E



Ein Buch ist nicht nur ein Freund, es schafft Dir neue Freunde.

Wenn Du ein Buch in Gedanken und im Geist besessen hast, bist Du bereichert.

Aber wenn Du das Buch weitergegeben hast, bist Du dreifach bereichert.

Henry Miller, The Books In My Life (1969)

REFERENCES

Ahmad, M.M., Roberts, J.N.J., Hardiman, E.M.E., Singh, R.R., Eltis, L.D.L. and Bugg, T.D.H.T. (2011) Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase. Biochemistry **50**, 5096–5107.

Alberts, B. (2008) Molecular Biology of the Cell. Garland Pub.

Alfieri, A., Malito, E., Orru, R., Fraaije, M.W. and Mattevi, A. (2008) Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. Proceedings of the National Academy of Sciences of the United States of America **105**, 6572–6577.

Alphand, V., Fraaije, M.W., Mihovilovic, M.D. and Ottolina, G. (2009) Second generation Baeyer–Villiger biocatalysts, in: Anthonsen, T., Fessner, W.-D. (Eds.), *Applied Biocatalysis. Stereoselective and Environmentally Friendly Reactions Catalysed by Enzymes*. Wiley-VCH, pp. 339–368.

Andreini, C., Bertini, I., Cavallaro, G., Holliday, G.L. and Thornton, J.M. (2008) Metal ions in biological catalysis: from enzyme databases to general principles. Journal of Biological Inorganic Chemistry **13**, 1205–1218.

Argos, P. (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. Journal of Molecular Biology **211**, 943–958.

Baeyer, A. and Villiger, V. (1899) Einwirkung des Caro'schen Reagens auf Ketone. Berichte der Deutschen Chemischen Gesellschaft **32**, 3625–3633.

Baldwin, T.O. and Ziegler, M.M. (1992) The biochemistry and molecular biology of bacterial bioluminescence, in: Müller, F. (Ed.), *Chemistry and Biochemistry of Flavoenzymes*. CRC Critical Reviews in Biochemistry, CRC Press, Boca Raton, FL, pp. 467–530.

Balke, K., Kadow, M., Mallin, H., Saß, S., Bornscheuer, U.T. (2012) Discovery , application and protein engineering of Baeyer-Villiger monooxygenases for organic synthesis. Organic & Biomolecular Chemistry **10**, 6249-6265.

Blom, C. (2000) Peroxide Chemistry. Wiley-VCH, Weinheim.

Bocola, M., Schulz, F., Leca, F., Vogel, A., Fraaije, M.W. and Reetz, M.T. (2005) Converting phenylacetone monooxygenase into phenylcyclohexanone monooxygenase by rational design: towards practical Baeyer-Villiger monooxygenases. Advanced Synthesis & Catalysis **347**, 979–986.

Bosello, M.M., Mielcarek, A.A., Giessen, T.W.T. and Marahiel, M.A.M. (2012) An enzymatic pathway for the biosynthesis of the formylhydroxyornithine required for rhodochelin iron coordination. Biochemistry **51**, 3059–3066.

Brink, ten, G.J., Arends, I.W.C.E. and Sheldon, R.A. (2004) The Baeyer−Villiger reaction: new developments toward greener procedures. Chemical Reviews. **104**, 4105–4124.

Chen, G., Kayser, M.M., Mihovilovic, M.D., Mrstik, M.E., Martinez, C.A. and Stewart, J.D. (1999) Asymmetric oxidations at sulfur catalyzed by engineered strains that overexpress cyclohexanone monooxygenase. New Journal of Chemistry **23**, 827–832.

Choi, H.S., Kim, J.K., Cho, E.H., Kim, Y.C., Kim, J.I. and Kim, S.W. (2003) A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. Biochemical and Biophysical Research Communications **306**, 930–936.

Cirino, P.C. and Arnold, F.H. (2002) Protein engineering of oxygenases for biocatalysis. Current Opinion in Chemical Biology **6**, 130–135.

Colonna, S., Gaggero, N., Carrea, G., Ottolina, G., Pasta, P. and Zambianchi, F. (2002) First asymmetric epoxidation catalysed by cyclohexanone monooxygenase. Tetrahedron letters **43**, 1797–1799.

Colonna, S., Pironti, V., Pasta, P. and Zambianchi, F. (2003) Oxidation of amines catalyzed by cyclohexanone monooxygenase. Tetrahedron letters **44**, 869–871.

Costas, A.M., White, A.K. and Metcalf, W.W. (2001) Purification and characterization of a novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. Journal of Biological Chemistry **276**, 17429–17436.

de Gonzalo, G., Mihovilovic, M.D. and Fraaije, M.W. (2010) Recent developments in the application of Baeyer-Villiger monooxygenases as biocatalysts. ChemBioChem **11**, 2208–2231.

de Jong, E., van Berkel, W.J., van der Zwan, R.P. and de Bont, J.A. (1992) Purification and characterization of vanillyl-alcohol oxidase from *Penicillium simplicissimum*. A novel aromatic alcohol oxidase containing covalently bound FAD. European Journal of Biochemistry **208**, 651–657.

Dixon, M. and Webb, E.C. (1979) Enzyme inhibition and activation. Enzymes Academic Press **332**.

Dolphin, C.T., Janmohamed, A., Smith, R.L., Shephard, E.A. and Phillips, I.R. (1997) Missense mutation in flavin-containing monooxygenase 3 gene, FMO3, underlies fish-odour syndrome. Nature Genetics **17**, 491–494.

Donoghue, N.A., Norris, D.B. and Trudgill, P.W. (1976) The purification and properties of cyclohexanone oxygenase from *Nocardia globerula* CL1 and *Acinetobacter* NCIB 9871. European Journal of Biochemistry **63**, 175–192.

Dudek, H.M., de Gonzalo, G., Torres Pazmiño, D.E., Stepniak, P., Wyrwicz, L.S., Rychlewski, L. and Fraaije, M.W. (2011) Mapping the substrate binding Sste of phenylacetone monooxygenase from *Thermobifida fusca* by mutational analysis. Applied and Environmental Microbiology **77**, 5730–5738.

Dudek, H.M., Torres Pazmiño, D.E., Rodríguez, C., Gonzalo, G., Gotor, V. and Fraaije, M.W. (2010) Investigating the coenzyme specificity of phenylacetone monooxygenase from *Thermobifida fusca*. Applied Microbiology and Biotechnology **88**, 1135–1143.

Entsch, B. and van Berkel, W.J. (1995) Structure and mechanism of para-hydroxybenzoate hydroxylase. FASEB Journal **9**, 476–483.

Evers, T.H., van Dongen, E.M.W.M., Faesen, A.C., Meijer, E.W. and Merkx, M. (2006) Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. Biochemistry **45**, 13183–13192.

Fink, M.J., Rial, D.V., Kapitánová, P., Lengar, A., Rehdorf, J., Cheng, Q., Rudroff, F. and Mihovilovic, M.D. (2012) Quantitative comparison of chiral catalysts selectivity and performance: a generic concept illustrated with cyclododecanone monooxygenase as Baeyer-Villiger biocatalyst. Advanced Synthesis & Catalysis **354**, 3491–3500.

Fischer, J.D., Holliday, G.L. and Thornton, J.M. (2010) The Cofactor database: organic cofactors in enzyme catalysis. Bioinformatics **26**, 2496–2497.

Forneris, F., Orru, R., Bonivento, D., Chiarelli, L.R. and Mattevi, A. (2009) ThermoFAD, a Thermofluor®-adapted flavin ad hocdetection system for protein folding and ligand binding. FEBS Journal **276**, 2833–2840.

Forney, F.W. and Markovetz, A.J. (1969) An enzyme system for aliphatic methyl ketone oxidation. Biochemical and Biophysical Research Communications **37**, 31–38.

Fraaije, M.W. (2003) The prodrug activator EtaA from *Mycobacterium tuberculosis* Is a Baeyer-Villiger monooxygenase. Journal of Biological Chemistry **279**, 3354–3360.

Fraaije, M.W., Kamerbeek, N.M., van Berkel, W.J. and Janssen, D.B. (2002) Identification of a Baeyer–Villiger monooxygenase sequence motif. FEBS letters **518**, 43–47.

Fraaije, M.W., Wu, J., Heuts, D.P.H.M., van Hellemond, E.W., Spelberg, J.H.L. and Janssen, D.B. (2005) Discovery of a thermostable Baeyer-Villiger monooxygenase by genome mining. Applied Microbiology and Biotechnology **66**, 393–400.

Franceschini, S.S., Fedkenheuer, M.M., Vogelaar, N.J.N., Robinson, H.H.H., Sobrado, P.P. and Mattevi, A.A. (2012) Structural insight into the mechanism of oxygen activation and substrate selectivity of flavin-dependent N-hydroxylating monooxygenases. Biochemistry **51**, 7043–7045.

Fruk, L., Kuo, C.-H., Torres, E. and Niemeyer, C.M. (2009) Apoenzyme reconstitution as a chemical tool for structural enzymology and biotechnology. Angewandte Chemie International Edition **48**, 1550–1574.

Ghisla, S. and Edmondson, D.E. (2009) Flavin Coenzymes, eLS. eLS, Chichester, UK.

Ghisla, S. and Massey, V. (1989) Mechanisms of flavoprotein-catalyzed reactions. European Journal of Biochemistry **181**, 1–17.

Gonzalo, G. de, Pazmiño, D.E.T., Ottolina, G., Fraaije, M.W. and Carrea, G. (2005) Oxidations catalyzed by phenylacetone monooxygenase from *Thermobifida fusca*. Tetrahedron: Asymmetry **16**, 3077–3083.

Griffin, M. and Trudgill, P.W. (1976) Purification and properties of cyclopentanone oxygenase of *Pseudomonas* NCIB 9872. European Journal of Biochemistry **63**, 199–209.

Harden, A. and Young, W.J. (1906) JSTOR: Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character, **77**, No. 519 (Apr. 12, 1906), pp. 405-420.

Heuts, D.P.H.M., Scrutton, N.S., McIntire, W.S. and Fraaije, M.W. (2009) What's in a covalent bond? FEBS Journal **276**, 3405–3427.

Hines, R.N., Hopp, K.A., Franco, J., Saeian, K. and Begun, F.P. (2002) Alternative processing of the human FMO6 gene renders transcripts incapable of encoding a functional flavin-containing monooxygenase. Molecular pharmacology **62**, 320–325.

Hollmann, F., Hofstetter, K. and Schmid, A. (2006) Non-enzymatic regeneration of nicotinamide and flavin cofactors for monooxygenase catalysis. Trends in Biotechnology **24**, 163–171.

Hollmann, F., Taglieber, A., Schulz, F. and Reetz, M.T. (2007) A light-driven stereoselective biocatalytic oxidation. Angewandte Chemie International Edition 46, 2903–2906.

Iwaki, H., Hasegawa, Y., Wang, S., Kayser, M.M. and Lau, P.C. (2002) Cloning and characterization of a gene cluster involved in cyclopentanol metabolism in *Comamonas sp*. strain NCIMB 9872 and biotransformations effected by *Escherichia coli*-expressed cyclopentanone 1, 2-monooxygenase. Applied and Environmental Microbiology **68**, 5671–5684.

Iwaki, H., Wang, S., Grosse, S., Lertvorachon, J. Yang, J., Konishi, Y., Hasegawa, Y., Lau, P.C.K. (2006) Pseudomonad cyclopentadecanone monooxygenase displaying an uncommon spectrum of Baeyer-Villiger oxidations of cyclic ketones. Applied and Environmental Microbiology **72,** 2707 - 2720.

Jensen, C.N., Cartwright, J., Ward, J., Hart, S., Turkenburg, J.P., Ali, S.T., Allen, M.J. and Grogan, G. (2012) A Flavoprotein monooxygenase that catalyses a Baeyer-Villiger reaction and thioether oxidation using NADH as the nicotinamide cofactor. ChemBioChem **13**, 872–878.

Johannes, T.W., Woodyer, R.D. and Zhao, H. (2005) Directed evolution of a thermostable phosphite dehydrogenase for NAD(P)H regeneration. Applied and Environmental Microbiology **71**, 5728–5734.

Johannes, T.W., Woodyer, R.D. and Zhao, H. (2006) Efficient regeneration of NADPH using an engineered phosphite dehydrogenase. Biotechnology and Bioengineering **96**, 18–26.

Kamerbeek, N.M., Fraaije, M.W. and Janssen, D.B. (2004) Identifying determinants of NADPH specificity in Baeyer-Villiger monooxygenases. European Journal of Biochemistry **271**, 2107–2116.

Kamerbeek, N.M., Janssen, D.B., van Berkel, W.J.H. and Fraaije, M.W. (2003) Baeyer–Villiger monooxygenases, an emerging family of flavin-dependent biocatalysts. Advanced Synthesis & Catalysis **345**, 667–678.

Kamerbeek, N.M., Moonen, M., Ven, J., Berkel, W., van Berkel, W.J.H., Fraaije, M.W. and Janssen, D.B. (2001) 4-Hydroxyacetophenone monooxygenase from Pseudomonas fluorescens ACB. A novel flavoprotein catalyzing Baeyer-Villiger oxidation of aromatic compounds. European Journal of Biochemistry **268**, 2547-2557.

Kamerbeek, N.M., Olsthoorn, A.J.J., Fraaije, M.W. and Janssen, D.B. (2003) Substrate specificity and enantioselectivity of 4-hydroxyacetophenone monooxygenase. Applied and Environmental Microbiology **69**, 419–426.

Kamerbeek, N.M., van der Ploeg, R., Olsthoorn, A.J.J., Tahallah, N., Heck, A.J., Malito, E., Janssen, D.B. and Fraaije, M.W. (2005) Exploring the role of the N-terminus of 4-hydroxyacetophenone monooxygenase, in: Nishino, T., Miura, R., Tanokura, M., Fukui, K. (Eds.), *Flavins and Flavoproteins 2005 : Proceedings of the 15. International Symposium.* ARchiTect Inc, Tokyo, pp. 173–178.

Kayser, M.M. (2009) "Designer reagents" recombinant microorganisms: new and powerful tools for organic synthesis. Tetrahedron **65**, 947–974.

Keller, S., Wage, T., Hohaus, K., Hölzer, M., Eichhorn, E. and van Pee, K.-H. (2000) Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens*. Angewandte Chemie International Edition **39**, 2300–2302.

Kopacz, M., Hollmann, F. and Fraaije, M.W. (2012) Enzyme engineering by cofactor redesign, in: Lutz, S., Bornscheuer, U.T. (Eds.), *Protein Engineering Handbook.* Vol. 3 Wiley-VCH, Weinheim, pp. 163–192.

Latham, J.A. and Walsh, C. (1986) Retention of configuration in oxidation of a chiral boronic acid by the flavoenzyme cyclohexanone oxygenase. Journal of the Chemical Society, Chemical Communications 527–528.

Latham, J.A., Branchaud, B.P., Chen, Y.C.J. and Walsh, C. (1986) Allylic and propargylic phenyl selenide oxygenation by cyclohexanone oxygenase: [2,3]-sigmatraopic rearrangement of the enzyme-generated selenoxide. Journal of the Chemical Society, Chemical Communications 528–530.

Lee, W.-H., Park, J.-B., Park, K., Kim, M.-D. and Seo, J.-H. (2007) Enhanced production of epsilon-caprolactone by overexpression of NADPH-regenerating glucose 6-phosphate dehydrogenase in recombinant *Escherichia coli* harboring cyclohexanone monooxygenase gene. Applied Microbiology and Biotechnology **76**, 329–338.

Letunic, I. and Bork, P. (2011) Interactive tree of life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Research **39**, W475–W478.

Li, L., Liu, X., Yang, W., Xu, F., Wang, W., Feng, L., Bartlam, M., Wang, L. and Rao, Z. (2008) Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: unveiling the long-chain alkane hydroxylase. Journal of Molecular Biology **376**, 453–465.

Malito, E., Alfieri, A., Fraaije, M.W. and Mattevi, A. (2004) Crystal structure of a Baeyer-Villiger monooxygenase. Proceedings of the National Academy of Sciences of the United States of America **101**, 13157–13162.

Massey, V. (1994) Activation of molecular oxygen by flavins and flavoproteins. Journal of Biological Chemistry **269**, 22459–22462.

Massey, V. and Palmer, G. (1966) On the existence of spectrally distinct classes of flavoprotein semiquinones. A new method for the quantitative production of flavoprotein semiquinones. Biochemistry **5**, 3181–3189.

McLeod, M.P., Warren, R.L., Hsiao, W.W.L., Araki, N., Myhre, M., Fernandes, C., Miyazawa, D., Wong, W., Lillquist, A.L., Wang, D., Dosanjh, M., Hara, H., Petrescu, A., Morin, R.D., Yang, G., Stott, J.M., Schein, J.E., Shin, H., Smailus, D., Siddiqui, A.S., Marra, M.A., Jones, S.J.M., Holt, R., Brinkman, F.S.L., Miyauchi, K., Fukuda, M., Davies, J.E., Mohn, W.W. and Eltis, L.D. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proceedings of the National Academy of Sciences of the United States of America **103**, 15582–15587.

Mewies, M., McIntire, W.S. and Scrutton, N.S. (1998) Covalent attachment of flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) to enzymes: the current state of affairs. Protein Science **7**, 7–20.

Mihovilovic, M.D. (2006) Enzyme Mediated Baeyer-Villiger Oxidations. Current Organic Chemistry **10**, 1265–1278.

Mihovilovic, M.D., Kapitán, P. and Kapitánová, P. (2008) Regiodivergent Baeyer–Villiger oxidation of fused ketones by recombinant whole-cell biocatalysts. ChemSusChem **1**, 143–148.

Mihovilovic, M.D., Müller, B. and Stanetty, P. (2002) Monooxygenase-mediated Baeyer− Villiger oxidations. European Journal of Organic Chemistry **2002**, 3711–3730.

Mihovilovic, M.D., Rudroff, F. and Grotzl, B. (2004) Enantioselective Baeyer-Villiger oxidations. Current Organic Chemistry **8**, 1057–1069.

Mirza, I.A., Yachnin, B.J., Wang, S., Grosse, S., Bergeron, H., Imura, A., Iwaki, H., Hasegawa, Y., Lau, P.C.K. and Berghuis, A.M. (2009) Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor. Journal of the American Chemical Society **131**, 8848–8854.

Miyamoto, M., Matsumoto, J., Iwaya, T., Itagaki, E. (1995) Bacterial steroid monooxygenase catalyzing the Baeyer-Villiger oxidation of C21-ketosteroids from *Rhodococcus rhodochrous*: the isolation and characterization. Biochimica et Biophysica Acta **1251**, 115-124.

Montersino, S. (2012) Strucutral and biochemical characterization of 3-hydroxybenzoate 6-hydroxylase. Thesis at Wageningen University

Montersino, S., Tischler, D., Gassner, G.T. and van Berkel, W.J.H. (2011) Catalytic and structural features of flavoprotein hydroxylases and epoxidases. Adv. Synth. Catal. **353**, 2301–2319.

Montersino, S.S. and van Berkel, W.J.H.W. (2012) Functional annotation and characterization of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1. Biochimica et Biophysica Acta - Bioenergetics **1824**, 433–442.

Nielsen, M., Lundegaard, C., Lund, O. and Petersen, T.N. (2010) CPHmodels-3.0 – remote homology modeling using structure-guided sequence profiles. Nucleic Acids Research **38**, W576–81.

Olucha, J.J. and Lamb, A.L.A. (2011) Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases. Bioorganic Chemistry **39**, 171–177.

Orru, R., Dudek, H.M., Martinoli, C., Torres Pazmiño, D.E., Royant, A., Weik, M., Fraaije, M.W. and Mattevi, A. (2011) Snapshots of enzymatic Baeyer-Villiger catalysis: oxygen activation and intermediate stabilization. Journal of Biological Chemistry **286**, 29284–29291.

Ottolina, G., Bianchi, S., Belloni, B., Carrea, G. and Danieli, B. (1999) First asymmetric oxidation of tertiary amines by cyclohexanone monooxygenase. Tetrahedron letters **40**, 8483–8486.

Panke, S., Witholt, B., Schmid, A. and Wubbolts, M.G. (1998) Towards a biocatalyst for (S)-styrene oxide production: characterization of the styrene degradation pathway of *Pseudomonas* sp. strain VLB120. Applied and Environmental Microbiology **64**, 2032–2043.

Peretz, M., Bogin, O., Tel-Or, S., Cohen, A., Li, G., Chen, J.S. and Burstein, Y. (1997) Molecular cloning, nucleotide sequencing, and expression of genes encoding alcohol dehydrogenases from the thermophile *Thermoanaerobacter brockii* and the mesophile *Clostridium beijerinckii*. Anaerobe **3**, 12–12.

Prieto, M.A. and Garcia, J.L. (1994) Molecular characterization of 4-hydroxyphenylacetate 3-hydroxylase of *Escherichia coli*. A two-protein component enzyme. Journal of Biological Chemistry **269**, 22823–22829.

Reiss, P.D., Zuurendonk, P.F. and Veech, R.L. (1984) Measurement of tissue purine, pyrimidine, and other nucleotides by radial compression high-performance liquid chromatography. Analytical Biochemistry **140**, 162–171.

Renz, M. and Meunier, B. (1999) 100 years of Baeyer-Villiger oxidations. European Journal of Organic Chemistry **1999**, 737–750.

Rial, D.V., Bianchi, D.A., Kapitanova, P., Lengar, A., van Beilen, J.B. and Mihovilovic, M.D. (2008) Stereoselective desymmetrizations by recombinant whole cells expressing the Baeyer–Villiger monooxygenase from *Xanthobacter* sp. ZL5: a new biocatalyst accepting structurally demanding substrates. European Journal of Organic Chemistry 1203–1213.

Riebel, A., de Gonzalo, G. and Fraaije, M.W. (2013) Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. Journal of Molecular Catalysis. B, Enzymatic **88**, 20–25.

Riebel, A., Dudek, H.M., Gonzalo, G., Stepniak, P., Rychlewski, L. and Fraaije, M.W. (2012) Expanding the set of rhodococcal Baeyer–Villiger monooxygenases by high-throughput cloning, expression and substrate screening. Applied Microbiology and Biotechnology **95**, 1479–1489.

Rioz-Martínez, A., Kopacz, M., de Gonzalo, G., Torres Pazmiño, D.E., Gotor, V. and Fraaije, M.W. (2011) Exploring the biocatalytic scope of a bacterial flavin-containing monooxygenase. Organic & Biomolecular Chemistry **9**, 1337–1341.

Roberts, S.M. and Wan, P.W. (1998) Enzyme-catalysed Baeyer–Villiger oxidations. Journal of Molecular Catalysis. B, Enzymatic **4**, 111–136.

Robinson, C.R. and Sauer, R.T. (1998) Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. Proceedings of the National Academy of Sciences of the United States of America **95**, 5929–5934.

Rudroff, F., Rydz, J., Ogink, F.H., Fink, M., Mihovilovic, M.D. (2007) Comparing the stereoselective biooxidation of cyclobutanones by recombinant strains expressing bacterial Baeyer–Villiger monooxygenases. Advanced Synthesis & Catalysis **349**, 1436–1444.

Ryerson, C.C., Ballou, D.P. and Walsh, C. (1982) Mechanistic studies on cyclohexanone oxygenase. Biochemistry **21**, 2644–2655.

Saheki, S., Takeda, A. and Shimazu, T. (1985) Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. Analytical Biochemistry **148**, 277–281.

Schafer, F.Q. and Buettner, G.R. (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Radical Biology & Medicine **30**, 1191–1212.

Schlaich, N.L. (2007) Flavin-containing monooxygenases in plants: looking beyond detox. Trends in Plant Science **12**, 412–418.

Schulz, F., Leca, F., Hollmann, F. and Reetz, M.T. (2005) Towards practical biocatalytic Baeyer-Villiger reactions: applying a thermostable enzyme in the gram-scale synthesis of optically-active lactones in a two-liquid-phase system. Beilstein Journal of Organic Chemistry **1**, 10–10.

Secundo, F., Fialà, S., Fraaije, M.W., de Gonzalo, G., Meli, M., Zambianchi, F. and Ottolina, G. (2010) Effects of water miscible organic solvents on the activity and conformation of the Baeyer-Villiger monooxygenases from *Thermobifida fusca* and *Acinetobacter calcoaceticus*: A comparative study. Biotechnology and Bioengineering **108**, 491–499.

Seto, M., Kimbara, K., Shimura, M., Hatta, T., Fukuda, M. and Yano, K. (1995) A novel transformation of polychlorinated biphenyls by *Rhodococcus sp*. strain RHA1. Applied and Environmental Microbiology **61**, 3353–3358.

Seto, M., Masai, E., Ida, M., Hatta, T., Kimbara, K., Fukuda, M. and Yano, K. (1995) Multiple polychlorinated biphenyl transformation systems in the gram-positive Bacterium *Rhodococcus* sp. strain RHA1. Applied and Environmental Microbiology **61**, 4510–4513.

Sheng, D., Ballou, D.P. and Massey, V. (2001) Mechanistic studies of cyclohexanone monooxygenase: Chemical properties of intermediates involved in catalysis. Biochemistry **40**, 11156–11167.

Stehr, M., Diekmann, H., Smau, L., Seth, O., Ghisla, S., Singh, M. and Macheroux, P. (1998) A hydrophobic sequence motif common to N-hydroxylating enzymes. Trends in Biochemical Sciences **23**, 56–57.

Stewart, J.D. (1997) A chemist's perspective on the use of genetically engineered microbes as reagents for organic synthesis. Biotechnology and Genetic Engineering Reviews **14**, 67–143.

Stewart, J.D. (1998) Cyclohexanone monpoxygenase: A useful reagent for asymmetric Baeyer-Villiger reactions. Current Organic Chemistry **2**, 195–216.

Stewart, J.D., Reed, K.W. and Kayser, M.M. (1996) "Designer yeast": a new reagent for enantioselective Baeyer–Villiger oxidations. Journal of the Chemical Society, Perkin Transactions 1 **8**, 755-757.

Strukul, G. (1998) Transition metal catalysis in the Baeyer–Villiger oxidation of ketones. Angewandte Chemie International Edition **37**, 1198–1209.

Szolkowy, C., Eltis, L.D., Bruce, N.C. and Grogan, G. (2009) Insights into sequence-activity relationships amongst Baeyer-Villiger monooxygenases as revealed by the intragenomic complement of enzymes from *Rhodococcus jostii* RHA1. ChemBioChem **10**, 1208–1217.

Tamura, K.K., Peterson, D.D., Peterson, N.N., Stecher, G.G., Nei, M.M. and Kumar, S.S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution **28**, 2731–2739.

Taylor, D.G. and Trudgill, P.W. (1986) Camphor revisited: studies of 2,5-diketocamphane 1,2-monooxygenase from *Pseudomonas putida* ATCC 17453. Journal of bacteriology **165**, 489–497.

Thapa, A., Shahnawaz, M., Karki, P., Raj Dahal, G., Sharoar, M.G., Yub Shin, S., Sup Lee, J., Cho, B. and Park, I.-S. (2008) Purification of inclusion body-forming peptides and proteins in soluble form by fusion to *Escherichia coli* thermostable proteins. BioTechniques **44**, 787–796.

Theorell, H. (1935) Das gelbe oxydationsferment. Biochemische Zeitschrift **278**, 263-290.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**, 4673–4680.

Tishkov, V.I. and Popov, V.O. (2006) Protein engineering of formate dehydrogenase. Biomolecular Engineering **23**, 22–22.

Tishkov, V.I., Galkin, A.G., Fedorchuk, V.V., Savitsky, P.A., Rojkova, A.M., Gieren, H. and Kula, M.R. (1999) Pilot scale production and isolation of recombinant NAD^+- and $NADP^+$-specific formate dehydrogenases. Biotechnology and Bioengineering **64**, 187–193.

Torres Pazmiño, D.E. and Fraaije, M.W. (2007) Discovery, redesign and applications of Baeyer-Villiger monooxygenases, in: Matsuda, T. (Ed.), *Future Directions in Biocatalysis*. Elsevier Sciences, pp. 107–128.

Torres Pazmiño, D.E., Snajdrova, R., Rial, D.V., Mihovilovic, M.D. and Fraaije, M.W. (2007) Altering the substrate specificity and enantioselectivity of phenylacetone monooxygenase by structure-inspired enzyme redesign. Advanced Synthesis & Catalysis **349**, 1361–1368.

Torres Pazmiño, D.E., Baas, B.J. and Janssen, D.B. (2008) Kinetic mechanism of phenylacetone monooxygenase from *Thermobifida fusca*. Biochemistry **47**, 4082–4093.

Torres Pazmiño, D.E., Snajdrova, R., Baas, B.-J., Ghobrial, M., Mihovilovic, M.D. and Fraaije, M.W. (2008) Self-sufficient Baeyer–Villiger monooxygenases: Effective coenzyme regeneration for biooxygenation by fusion engineering. Angewandte Chemie International Edition. **47**, 2275–2278.

Torres Pazmiño, D.E., Riebel, A., de Lange, J., Rudroff, F., Mihovilovic, M.D. and Fraaije, M.W. (2009) Efficient biooxidations catalyzed by a new generation of self-sufficient Baeyer-Villiger monooxygenases. ChemBioChem **10**, 2595–2598.

Torres Pazmiño, D.E., Dudek, H.M. and Fraaije, M.W. (2010) Baeyer-Villiger monooxygenases: recent advances and future challenges. Current Opinion in Chemical Biology **14**, 138–144.

Torres Pazmiño, D.E., Winkler, M., Glieder, A. and Fraaije, M.W. (2010) Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications. Journal of Biotechnology **146**, 9–24.

Turfitt, G.E. (1948) The microbiological degradation of steroids; fission of the steroid molecule. Biochemical Journal **42**, 376–383.

Valton, J., Mathevon, C., Fontecave, M., Nivière, V. and Ballou, D.P. (2008) Mechanism and regulation of the two-component FMN-dependent monooxygenase ActVA-ActVB from *Streptomyces coelicolor*. Journal of Biological Chemistry **283**, 10287–10296.

van Beilen, J.B., Mourlane, F., Seeger, M.A., Kovac, J., Li, Z., Smits, T.H.M., Fritsche, U. and Witholt, B. (2003) Cloning of Baeyer-Villiger monooxygenases from *Comamonas*, *Xanthobacter* and *Rhodococcus* using polymerase chain reaction with highly degenerate primers. Environmental Microbiology **5**, 174–182.

van Berkel, W.J.H., Kamerbeek, N.M. and Fraaije, M.W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. Journal of Biotechnology **124**, 670–689.

van Bloois, E., Torres Pazmiño, D.E., Winter, R.T. and Fraaije, M.W. (2010) A robust and extracellular heme-containing peroxidase from Thermobifida fusca as prototype of a bacterial peroxidase superfamily. Applied Microbiology and Biotechnology **86**, 1419–1430.

van den Heuvel, R.H., Tahallah, N., Kamerbeek, N.M., Fraaije, M.W., van Berkel, W.J., Janssen, D.B. and Heck, A.J. (2005) Coenzyme binding during catalysis is beneficial for the stability of 4-hydroxyacetophenone monooxygenase. Journal of Biological Chemistry **280**, 32115–32121.

van der Donk, W.A. and Zhao, H. (2003) Recent developments in pyridine nucleotide regeneration. Current Opinion in Biotechnology **14**, 421–426.

van der Geize, R. and Dijkhuizen, L. (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Current Opinion in Microbiology **7**, 255–261.

Van Der Werf, M.J. (2000) Purification and characterization of a Baeyer-Villiger mono-oxygenase from *Rhodococcus erythropolis* DCL14 involved in three different monocyclic monoterpene degradation pathways. Biochemical Journal **347 Pt 3**, 693–701.

Villalobos, A., Ness, J.E., Gustafsson, C., Minshull, J. and Govindarajan, S. (2006) Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformatics **7**, 285–293.

Viviani, V.R. (2002) The origin, diversity, and structure function relationships of insect luciferases. Cellular and molecular life sciences **59**, 1833–1850.

Vrtis, J.M.J., White, A.K.A., Metcalf, W.W.W. and van der Donk, W.A.W. (2002) Phosphite dehydrogenase: a versatile cofactor-regeneration enzyme. Angewandte Chemie International Edition **41**, 3257–3259.

Walsh, C. (1980) Flavin coenzymes: at the crossroads of biological redox chemistry. Accounts of Chemical Research.

Walsh, C.T. and Chen, Y.C.J. (1988) Enzymic Baeyer-Villiger oxidations by flavin-dependent monooxygenases. Angewandte Chemie International Edition **27**, 333–343.

Walsh, C.T. and Wencewicz, T.A. (2013) Flavoenzymes: versatile catalysts in biosynthetic pathways. Natural Product Reports **30**, 175–200.

Walsh, F.C. (1979) Enzymatic Reaction Mechanism. W H Freeman & Company.

Warburg, O. and Christian, W. (1933) Über das gelbe Ferment und seine Wirkungen. Biochemische Zeitschrift **266**, 377–411.

Warburg, O. and Christian, W. (1936) Pyridin, der wasserstoffübertragende Bestandteil von Gärungsfermenten. Helvetica Chimica Acta **19**, E79–E88.

Wichmann, R. and Vasic-Racki, D. (2005) Cofactor regeneration at the lab scale. Advances in Biochemical Engineering / Biotechnology **92**, 225–260.

Wierenga, R.K., Terpstra, P. and Hol, W.G. (1986) Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint. Journal of Molecular Biology **187**, 101–107.

Williams, R.E. and Bruce, N.C. (2002) "New uses for an Old Enzyme"– the Old Yellow enzyme family of flavoenzymes. Microbiology **148**, 1608-1614.

Woodyer, R., van der Donk, W.A. and Zhao, H. (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry **42**, 11604–11614.

Woodyer, R., van der Donk, W.A. and Zhao, H. (2006) Optimizing a biocatalyst for improved NAD(P)H regeneration: directed evolution of phosphite dehydrogenase. Combinatorial Chemistry & High Throughput Screening **9**, 237–245.

Woodyer, R., Zhao, H. and van der Donk, W.A. (2005) Mechanistic investigation of a highly active phosphite dehydrogenase mutant and its application for NADPH regeneration. FEBS Journal **272**, 3816–3827.

Wu, S., Acevedo, J.P. and Reetz, M.T. (2010) Induced allostery in the directed evolution of an enantioselective Baeyer-Villiger monooxygenase. Proceedings of the National Academy of Sciences **107**, 2775–2780.

Zambianchi, F., Pasta, P., Carrea, G., Colonna, S., Gaggero, N. and Woodley, J.M. (2002) Use of isolated cyclohexanone monooxygenase from recombinant *Escherichia coli* as a biocatalyst for Baeyer-Villiger and sulfide oxidations. Biotechnology and Bioengineering **78**, 489–496.

Ziegler, D.M. (1990) Flavin-containing monooxygenases: enzymes adapted for multisubstrate specificity. Trends in Pharmacological Sciences **11**, 321–324.

Die wirklich guten Dinge im Leben haben etwas Lässiges.

Sie liegen wie Kühe auf der Wiese.

Friedrich Nietzsche

NEDERLANDSE SAMENVATTING

GESCHIKT VOOR DE GEINTERESSEERDE LEEK

GELE VOORUITZICHTEN

EEN STUDIE NAAR FLAVOPROTEIN MONOOXYGENASES

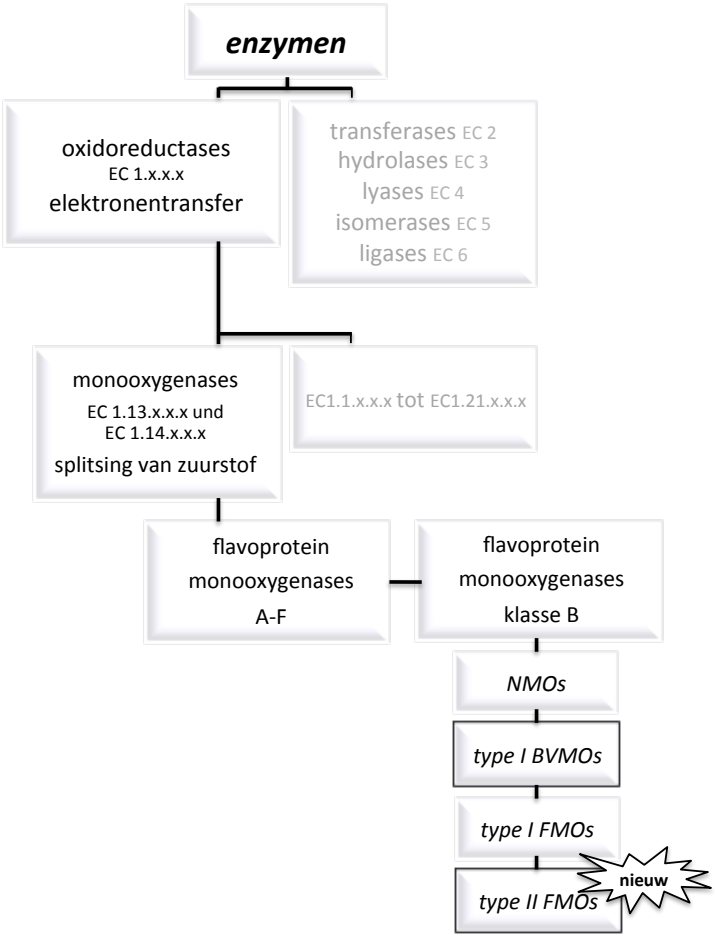
Enzymen, ook wel biokatalysatoren genoemd, zijn de werkers die in een levend organisme alles op-, af- en ombouwen. In elk orgaan en elke cel zijn veel verschillende enzymen aanwezig die de meest diverse taken kunnen uitvoeren. Zoals DNA het bouwplan van het leven genoemd wordt, kan men enzymen de huiselfjes van het leven noemen. Want niet alleen zijn ze onmisbaar in alle levende organismen, ze zijn ondertussen ook onmisbaar in ons dagelijkse leven. We gebruiken enzymen namelijk voor vele taken. Zo zitten er enzymen in wasmiddel en op teststrips waarmee kwaliteit van levensmiddelen getest kan worden. Dit zijn maar twee van de vele diverse voorbeelden die genoemd zouden kunnen worden.

Sinds de ontdekking van enzymen wordt er veel onderzoek naar gedaan. Niet alleen de werking in levende organismes wordt onderzocht, maar ook wordt er gekeken naar het gebruik van enzymen in de chemische/biotechnologische industrie. Zo worden enzymen tegenwoordig vaak gebruikt in milieuvriendelijkere en/of goedkopere alternatieven voor klassieke chemische industriële productie processen.

Enzymen zijn in staat om zeer complexe moleculen te produceren die via een traditionele chemische route alleen zeer moeilijk, en daarmee vaak duur, te maken zijn. Daarnaast worden enzymen ook steeds vaker gebruikt om alledaagse chemicaliën te produceren. Enzymen kunnen in staat zijn onder mildere omstandigheden producten te maken waarbij voor de klassieke chemische route bijvoorbeeld hoge druk of zeer giftige oplosmiddelen nodig zijn. Tevens zorgt de hoge specificiteit van enzymkatalyse, ook wel biokatalyse genoemd, ervoor dat producten vaak al zuiverder zijn waardoor er minder gecompliceerde zuiveringsstappen nodig zijn. Bovendien zijn enzymen niet alleen zelf van een biologische oorsprong (opgebouwd uit aminozuren), ze faciliteren ook het gebruik van hernieuwbare grondstoffen en afvalproducten – en produceren daarmee weer

biologisch afbreekbare producten. Dat alles komt het milieu ten goede en maakt enzymen ideaal gereedschap in de industrie.

De enzymen die in dit proefschrift worden behandeld (figuur NL-1), behoren allemaal tot de klasse oxidoreductases (EC1.x.x.x). Oxidoreductases zijn verantwoordelijk voor de overdracht van elektronen in biochemische reacties. Enzymen worden daarnaast nog verder geclassificeerd afhankelijk van de reactie de gekatalyseerd wordt. Onder oxidoreductases vallen weer 21 enzymsoorten. De enzymen die in dit proefschrift behandeld worden, behoren tot monooxygenases (EC 1.13.xx en EC 1.14.xx). Deze subklasse is in staat om zuurstof (O_2) te splitsen. Bij reacties van deze subklasse wordt daarna één van de zuurstofatomen in een organisch substraat ingebouwd terwijl het andere zuurstofatoom tot water (H_2O) gereduceerd wordt. Dissociatie van zuurstof is niet makkelijk en is alleen maar mogelijk door toevoeging van elektronen. Om elektronen toe te voegen gebruikt het enzym een “hulpgereedschap”, namelijk een cofactor. In de natuur bestaan vele verschillende “hulpgereedschappen” die op verschillende manieren eenzelfde chemische overdracht mogelijk maken. De enzymen die hier besproken worden, gebruiken als cofactor een flavine (zie paragraaf 1.2). Een flavine cofactor is een geel molecuul waardoor ook de enzymen die flavine bevatten geel zijn. Onder monooxygenases zijn er zes subklassen (A-F) die flavine als cofactor gebruiken. Deze subklassen worden daarom flavoprotein monooxygenases genoemd. Alhoewel deze enzymen dus allemaal flavine nodig hebben om zuurstof te activeren, katalyseren ze vervolgens verschillende reacties. Het zuurstofatoom wordt door verschillende enzymen in verschillende substraten op verschillende plekken ingebouwd.

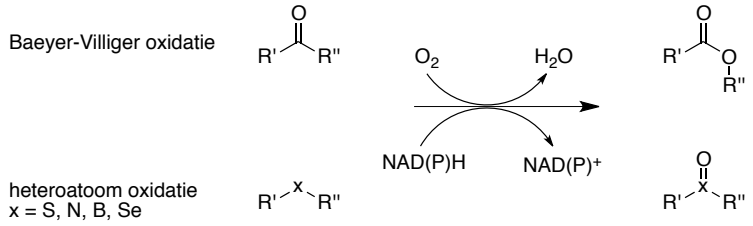


Figuur NL-1: Overzicht van de classificatie van de in dit proefschrift besproken enzymen.

In dit proefschrift worden alleen klasse B monooxygenases behandeld. Dit zijn enzymen die Baeyer-Villiger oxidaties en oxidaties van heteroatomen verrichten (figuur NL-2).

In figuur NL-2 is te zien dat er nog een ingrediënt nodig is om de reactie te laten verlopen, namelijk een coenzym in de vorm van NAD(P)H (zie paragraaf 1.3). Dit levert de elektronen die de flavine cofactor gebruikt om het zuurstof te activeren. Een bijzonderheid bij klasse B monooxygenases is dat ze zelf in staat zijn om de elektronen van

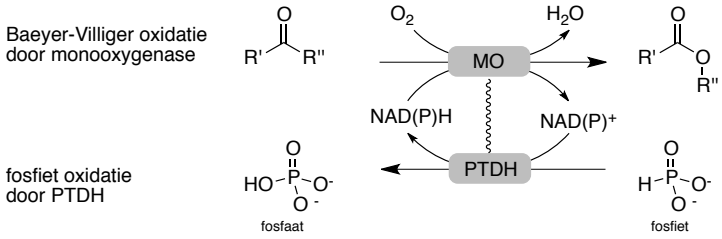
het coenzym NAD(P)H naar de flavine te transporteren. De enzymen uit andere enzyme klassen (C-F) hebben daarvoor nog weer een extra enzym nodig, een aparte reductase.



Figuur NL-2: Reacties die door klasse B monooxygenasen gekatalyseerd worden. R stelt verschillende organische reststukken van het molecuul voor en x staat voor zwavel, stikstof, boor of selenium.

Er bestaan twee versies van bovengenoemd coenzym met als enige verschil een extra fosfaatgroep; NADH en NADPH. De fosfaatgroep heeft geen invloed op de reactiviteit van het coenzym. Tot voor kort werd gedacht dat klasse B monooxygenases alleen met NADPH functioneren. Wij hebben echter een nieuwe subklasse van enzymen gevonden die ook NADH accepteert (Hoofdstuk 5 en 6). Het probleem van de NADH en NADPH coenzymen zijn echter de kosten ervan; NADH is al zeer duur, en NADPH is nog zo'n tien keer zo duur. Dit maakt het gebruik van NAD(P)H-afhankelijke enzymen minder aantrekkelijk voor industriële toepassingen. Daarom wordt er veel gedaan om het coenzym binnen de reactie te hergebruiken zodat er minder van het dure coenzym nodig is. De meest aantrekkelijke methode om dit te doen lijkt om nog een extra enzym aan de reactie toe te voegen. Dit extra enzym verbruikt een "offer"-substraat en zet daarbij het $NAD(P)^+$ weer om naar NAD(P)H. Door onze onderzoeksgroep is een expressiesysteem ontwikkeld dat op DNA niveau een regeneratie-enzym heeft gekoppeld aan een monooxygenase. Dit expressiesysteem zorgt ervoor dat ons eiwit geproduceerd wordt door de bacterie *Escherichia coli.* Het nieuwe expressiesysteem fuseert beide enzymen met een linker en zorgt zo voor de productie van een bifunctioneel enzym (figuur NL-3); het CRE-Systeem (*coenzyme regenerating enzyme*). Het enzym wat in dit geval aan de monooxygenase is gekoppeld is een fosfietdehydrogenase (PTDH). In dit systeem hoeft

men dus alleen wat goedkoop substraat voor het PTDH toe te voegen (fosfiet) zodat het systeem lang kan doorwerken met slechts een kleine hoeveelheid van het dure coenzym.



Figuur NL-3: Reacties die plaatsvinden met behulp van het bifunctionele CRE enzym. MO = monooxygenase, PTDH = fosfiet dehydrogenase, R geeft verschillende organische reststukken van het molecuul aan.

In Hoofdstuk 2 wordt de tweede generatie van het CRE-systeem beschreven (CRE2). Het originele systeem werd gelimiteerd door de lage stabiliteit van het PTDH die de looptijd van het bifunctionele systeem negatief beïnvloedde. Om dit te verbeteren hebben we gekozen voor een stabielere, genetisch geoptimaliseerde versie van het PTDH. Het gebruik van het nieuwe PTDH voor het CRE2-systeem heeft als bijkomend voordeel dat dit PTDH tot hoge expressie heeft. Dat wil zeggen een cel kan relatief makkelijk veel van dit eiwit produceren. Doordat het PTDH gefuseerd is aan een monooxygenase, verhoogt dit ook de expressie van het bifunctionele enzym. Daarnaast hebben we ook de zuivering van het bifunctionele enzym makkelijker gemaakt door er een zogenaamde His-*tag* aan te plakken. Dit nieuwe systeem is succesvol getest met andere NAD(P)H-afhankelijke enzymen en zeer efficiënt bevonden. Als laatste is ook nog onderzoek gedaan naar de sequentie van de linker die de beide enzymen aan elkaar koppelt. Hier is echter geen sequentie uitgekomen die het gehele systeem verbeterde.

Ondanks dat enzymen zeer veelzijdig zijn, is er nog niet voor elke reactie of substraat de passende biokatalysator gevonden. Daarom wordt er veel onderzoek gedaan naar het aanpassen van bekende enzymen om ze andere reacties te laten uitvoeren of andere substraten te accepteren. Helaas zijn de methoden hiervoor erg onnauwkeurig en voornamelijk gebaseerd op het *trial-and-error* principe. Deze methoden en experimenten zorgen echter wel voor meer informatie over enzymen en enzymklassen. Zo bestaan er

voor de verschillende flavoprotein monooxygenases speciale sequentie motieven (*fingerprint*-motief) waarmee potentiele kandidaten in databanken opgezocht kunnen worden. In de verschillende databanken zijn alle bekende genomen, al het DNA van een organisme, opgeslagen. In de databanken zijn ook vele bekende enzymsequenties (de volgorde van de aminozuren waaruit ze bestaan) en de genen (een stuk DNA wat voor een enzym codeert) te vinden.

Voor dit promotie-traject is het genoom van de bodembacterie *Rhodococcus jostii* RHA1 met behulp van *fingerprint*-motieven doorzocht en zijn interessante kandidaten geïdentificeerd. *R. jostii* herbergt een uitzonderlijk groot aantal oxidatieve enzymen waaronder meer dan 40 potentiele flavoprotein monooxygenases. De focus van dit promotieonderzoek lag op twee subgroepen van klasse B monooxygenases; Type I Baeyer-Villiger monooxygenases (BVMOs) en flavine bevattende monooxygenases (FMOs). Hoewel dergelijke enzymen niet ongewoon zijn in bacteriën en schimmels, heeft elk micro-organisme doorgaans maar ca. één BVMO gen en FMO's zijn nog zeldzamer. Met 40 potentiele kandidaten lijkt *R. jostii* dus een ongekend grote bron te zijn en vanwege de kennis over de capaciteiten van dit organisme zijn de verwachtingen voor deze kandidaten hoog. In totaal zijn er 23 BVMO's, zeven FMO's en nog verschillende andere flavine afhankelijke enzymen geïdentificeerd, gekloneerd in *E. coli* en tot expressie gebracht en geanalyseerd. In Hoofdstuk 3 worden de resultaten over de BVMOs besproken en in Hoofdstuk 4 & 5 de resultaten over de FMOs. Om het analyseren van een groot aantal samples makkelijker en efficiënter te maken zijn er twee nieuwe analysemethodes ontwikkeld. Met de eerste methode is het mogelijk de concentratie van NADPH-afhankelijke enzymen te bepalen in het cellysaat, dit is de inhoud van een cel nadat een cel is opengebroken. Dit zorgt ervoor dat het niet nodig is om complexe en tijdrovende zuiveringsstappen uit te voeren. De tweede methode is een kwantitatieve test voor het verbruik van NAD(P)H waarmee indirect enzymactiviteit bepaald wordt. Met deze twee methoden kan een enzym snel voor zeer veel verschillende substraten getest worden. De combinatie van beide testen staat een snelle, relatief goedkope en accurate analyse van nieuwe enzymen toe. Bovendien hebben we op grond van de grote

hoeveelheid data een nieuw sequentie motief, *fingerprint*-motief, voor de identificatie van BVMOs ontdekt.

Uit de resultaten van de FMO analyse (Hoofdstuk 4) bleek dat het hier ging om een nieuwe subgroep van klasse B monooxygenases; Type II FMOs. Alle zeven kandidaten hebben het typische FMO sequentie motief maar katalyseren de omzetting van typische substraten van Type I FMOs slecht. Ze zijn, in tegenstelling tot de andere subgroepen van klasse B monooxygenases, in staat om niet alleen NADPH, maar ook het goedkopere NADH als coenzym te gebruiken. Alleen al deze vaardigheid maakt ze zeer interessante biokatalysatoren waarvan het substraatspectrum dringend onderzocht moet worden. Bovendien waren drie enzymen van deze nieuwe subgroep in staat om Baeyer-Villiger oxidaties te katalyseren, een reactie die voorheen niet bij FMOs geobserveerd werd. Dit in combinatie met de coenzym flexibiliteit maakt deze enzymen heel aantrekkelijk. Hoofdstuk 5 gaat over het verdere onderzoek aan deze drie enzymen met betrekking tot substraatbereik en kinetische eigenschappen. Aan de hand van modellen (zie Figure 5-4) zijn de aminozuren, de bouwstenen van enzymen, van het potentiele actieve centrum geïdentificeerd en hun invloed op de enzymactiviteit door mutatiestudies bevestigd. Alle resultaten duiden op een zeer interessante nieuwe enzymklasse. Een bijzonder kenmerk van deze enzymklasse lijkt een extra lange sequentie aan het enzymeinde te zijn. Hiermee zijn meer potentiele kandidaten van deze enzymklasse in de verschillende databanken te vinden.

Naast de ontwikkeling van nieuwe methoden ter karakterisering van oxidatieve enzymen heeft dit promotieonderzoek voor een verhoogd aantal gekarakteriseerde BVMOs gezorgd. Maar dat niet alleen; door dit promotieonderzoek is er een nieuwe en veelbelovende subgroep van de flavoprotein monooxygenase familie ontdekt. Deze nieuwe Type II FMOs hebben de potentie de nieuwe sterren van de klasse B monooxygenasen te worden, met glanzende gele vooruitzichten…

Heute mach ich mir kein Abendbrot.

Heute mach ich mir Gedanken.

Wolfgang Neuss

Deutsche Zusammenfassung für den interessierten Laien

GELBE AUSSICHTEN

STUDIEN ZU FLAVOPROTEIN MONOOXYGENASEN

Enzyme, auch Biokatalysatoren genannt, sind sozusagen die Handwerker, die in lebenden Organismen alles auf-, ab- oder umbauen. Jedem Organ und jeder Zelle stehen die unterschiedlichsten Enzyme für die verschiedensten Aufgaben zur Verfügung. Während die DNA nur die Bauanleitung darstellt, sind die Enzyme durchaus als die Heinzelmännchen des Lebens zu betrachten. Und nicht nur das, sie sind auch in der Lage, ihre Tätigkeiten außerhalb vom lebenden Organismus durchzuführen, z.B. in der Waschmaschine oder auf Teststreifen in der Lebensmittelkontrolle.

Seit ihrer Entdeckung werden Enzyme ausführlich studiert, insbesondere für den Einsatz in der chemischen Industrie. Wurden sie zu Beginn eher als eine Spielerei der Natur betrachtet, so hat sich daraus mittlerweile eine eigene Disziplin entwickelt – die Biochemie und Biotechnologie – die sich derzeit zu einer ernstzunehmenden Alternative zur klassischen chemischen Industrie etabliert, insbesondere im Hinblick auf umweltfreundliche Verfahrensweisen.

Enzyme sind nicht nur in der Lage Stoffe herzustellen, die auf chemischem Wege nur sehr kompliziert und somit teuer zu produzieren sind, immer häufiger ersetzen sie auch Verfahren bei der Produktion von Alltags-Chemikalien. Durch ihre Fähigkeit z.B. unter milden Bedingungen und im Wässrigen zu arbeiten, machen sie Extrembedingungen (hoher Druck oder Temperatur), sowie toxische Lösungsmittel überflüssig. Aufgrund der hohen Spezifität der Enzymkatalyse ist oft auch die Reinheit der Produkte höher als im konventionellen Verfahren und auf teure und komplizierte Aufreinigungsmethoden kann verzichtet werden. Außerdem sind sie nicht nur selbst biologischer Herkunft (aufgebaut aus Aminosäuren), sie ermöglichen oft auch die Verwertung von nachwachsenden Rohstoffen und Abfallprodukten, und generieren ihrerseits wieder biologisch abbaubare Produkte. Das alles schont Umwelt und Geldbeutel und macht Enzyme zu idealen Werkzeugen für die Industrie.

Die Enzyme, die in dieser Dissertation behandelt werden (Abbildung D-1) gehören ganz allgemein zur Klasse der Oxidoreduktasen (EC1.x.x.x) und sind somit für den Transfer von Elektronen in biochemischen Reaktionen verantwortlich. Allerdings gibt es bei den Enzymen, wie im Handwerk unterschiedliche Berufe, d.h. typische Reaktionen, die von den einzelnen Enzymklassen katalysiert werden. Unter den Oxidoreduktasen gibt es derzeit 21 verschiedene Spezialisierungen und die Enzyme, die in dieser Dissertation erforscht werden, gehören zu den Monooxygenasen (EC 1.13.xx und EC 1.14.xx). Diese Subklasse ist in der Lage Sauerstoff (O_2) zu spalten. Bei den Reaktionen dieser Klasse wird daraufhin eines der Sauerstoffatome in ein organisches Substrat eingebaut, während das andere zu Wasser reduziert wird. Die Spaltung von Sauerstoff ist nicht einfach und nur durch die Zuführung von Elektronen zu bewerkstelligen. Das Enzym benötigt dazu als eine Art Werkzeug einen Cofaktor. In der Natur gibt es verschiedene solcher „Werkzeuge“, die letztlich alle auf unterschiedliche Weise dieselbe chemische Transaktion ermöglichen. Im Falle der hier betrachteten Enzyme ist der verwendete Cofaktor das Flavin (siehe Absatz 1.2), ein gelber Stoff, der auch den Enzymen die ihn beinhalten eine gelbe Farbe verleiht. Unter den Monooxygenasen gibt es 6 Subklassen (A – F) die als Cofaktor „das Werkzeug“ Flavin für die Aktivierung des Sauerstoff verwenden, sie werden als Flavoprotein Monooxygenasen bezeichnet. Allerdings katalysieren die Enzyme dieser Gruppe jedes eine leicht unterschiedliche Reaktion, d.h. sie fügen das Sauerstoffatom in jeweils unterschiedlichen Substraten an jeweils unterschiedlichen Stellen ein.

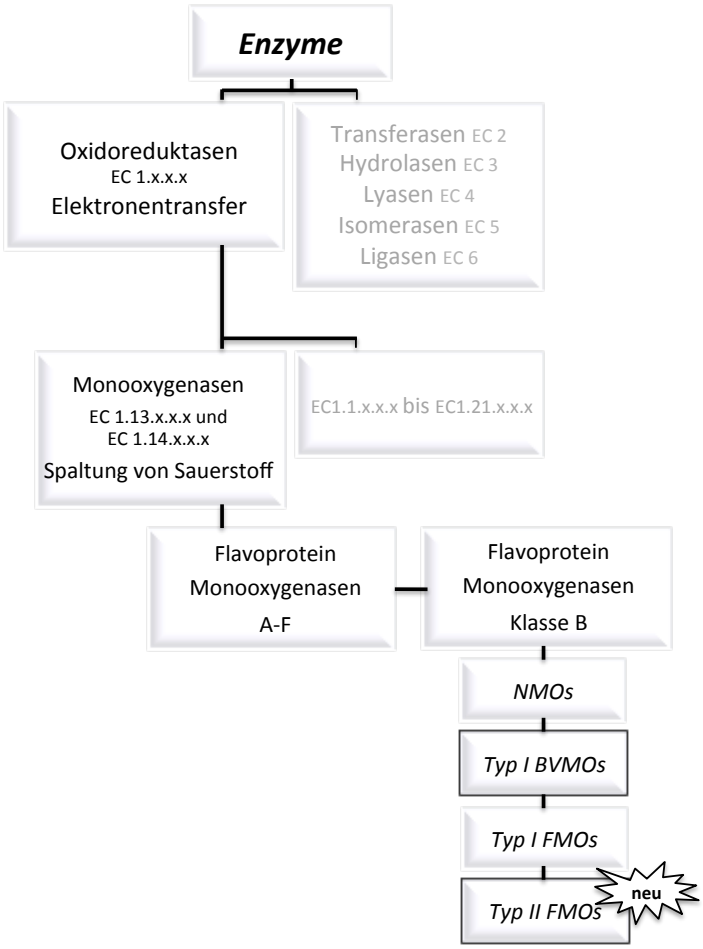


Abbildung D-1: Übersicht über die Klassifikation der hier diskutierten Enzyme.

In dieser Arbeit werden ausschließlich Enzyme der Subklasse B behandelt, Enzyme deren „Handwerk“ Baeyer-Villiger Oxidationen und die Oxidation von Heteroatomen darstellt (Abbildung D-2).

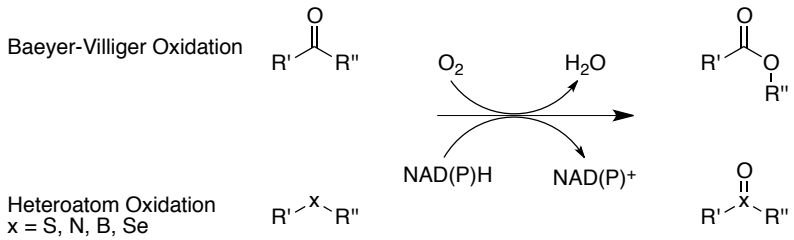


Abbildung D-2: Reaktionen katalysiert von Klasse B Monooxygenasen. R steht als Platzhalter für unterschiedliche organische Reste im Molekül, x steht als Platzhalter für eines der folgenden Elemente: Schwefel, Stickstoff, Bor, Selen.

Wie in Abbildung D-2 zu sehen ist, wird für die Reaktion außer Sauerstoff noch eine weitere Zutat benötigt, das Coenzym NAD(P)H (siehe Absatz 1.3). Dieses liefert die Elektronen, die dann vom Flavin auf den Sauerstoff übertragen werden können. Eine Besonderheit der Klasse B Monooxygenasen ist, dass sie selbst in der Lage sind, die Elektronen vom Coenzym NAD(P)H auf das Flavin zu transportieren. Die Monooxygenasen der anderen Klassen (C-F) benötigen dafür ein extra Enzym, eine separate Reduktase.

Das Coenzym existiert in zwei Versionen, die sich strukturell nur durch eine Phosphatgruppe unterscheiden (NADH und NADPH). Diese hat zwar keinen Einfluss auf die Funktion, aber die Klasse B Monooxygenasen funktionieren ausschließlich mit NADPH – das dachte man zumindest bisher: wir haben eine neue Subklasse gefunden, welche auch NADH akzeptiert, mehr dazu in **Kapitel 5** und **6**. Das Problem mit diesem Coenzym sind die Kosten: NADH ist schon teuer, aber NADPH ist ca. zehnmal teurer, was grundsätzlich alle Enzyme, die auf dieses Coenzym angewiesen sind, für die Industrie unattraktiv macht. Es wird daher viel Aufwand betrieben, um dieses Coenzym innerhalb eines Reaktionsansatzes zu recyceln. Als geeignetste Methode hat sich dabei die Verwendung eines weiteren Enzyms im Reaktionsansatz erwiesen, welches in einer zweiten Reaktion (unter Verbrauch eines billigen „Opfer"-Substrats) das verbrauchte $NAD(P)^+$ wieder zu NAD(P)H regeneriert. In unserem Labor haben wir ein Expressionssystem entwickelt, das auf DNA Ebene sozusagen ein solches Regenerationsenzym (in unserem Fall die Phosphitdehydrogenase, PTDH) fest mit einer

Monooxygenase verbindet und beide Enzyme zu einem bifunktionellen Enzym fusioniert (Abbildung D-3): das CRE-System (***C**oenzyme **R**egenerating **E**nzyme*). Mit diesem System erhält man einen einfach zu handhabenden Biokatalysator, der unter Zugabe des günstigen Opfer-Substrates Phosphit mit einer sehr geringen Menge des teuren Coenzymes auskommt.

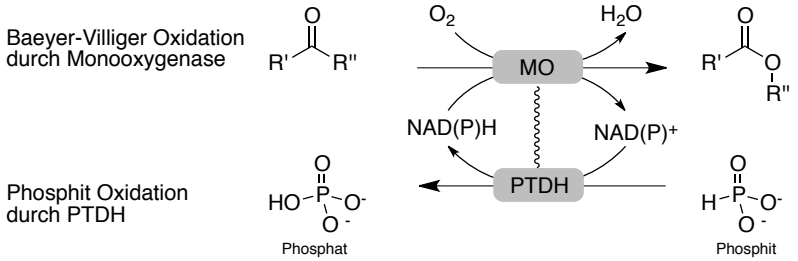


Abbildung D-3: Reaktionen des bifunktionellen CRE Enzyms. MO = Monooxygenase, PTDH = Phosphit Dehydrogenase, R steht als Platzhalter für unterschiedliche organische Reste im Molekül.

In **Kapitel 2** wird die nächste Generation dieses Systems beschrieben (CRE2). Ein limitierender Faktor im originalen System war die geringe Stabilität der PTDH, welche die Laufzeit des Biokatalysators beschränkte. Zur Optimierung haben wir eine stabilere und genetisch optimierte Version der PTDH verwendet. Da die PTDH ein leicht zu exprimierendes Enyzm ist, wirkt sie bei der Expression des bifunktionalen Enzyms nun als eine Art Zugpferd und erhöht somit zusätzlich die Expression der Monooxygenase. Als weitere Neuerung haben wir einen kleinen Anhang (His-*tag*) angefügt, der die Aufreinigung des Biokatalysators erleichtert. Das neue System wurde erfolgreich auch mit anderen NAD(P)H abhängigen Enzymen getestet und hat sich als äußerst effizient erwiesen. Des Weiteren wurde die Sequenz, welche die beiden Enzyme miteinander verbindet (linker) genauer untersucht, es konnte aber keine alternative Sequenz gefunden werden welche die Eigenschaften des Gesamtkonstrukts verbesserte.

So vielseitig Enzyme auch sind, so ist doch noch nicht für alle Reaktionen und Substrate das passende Werkzeug, sprich der optimale Biokatalysator gefunden. Derzeit wird viel daran geforscht, um einzelne, bekannte Enzyme so zu verändern, dass sie auch alternative Reaktionen durchführen oder andere Substrate akzeptieren. Diese Verfahren sind oft

noch sehr unpräzise und laufen meist nach dem *Trial-and-Error* Prinzip ab. Dennoch ermöglichen solche Experimente genauere Informationen über die einzelnen Enzyme und Enzymklassen. So existiert z.B. für die verschiedenen Untergruppen der Flavoprotein Monooxygenasen ein spezielles Sequenzmotiv (*fingerprint*–Motiv), anhand dessen potentielle Kandidaten in Datenbanken aufgespürt werden können, die sämtliche analysierten DNA- und Enzymsequenzen sammeln.

Für diese Dissertation wurde das Genom des Bodenbakteriums *Rhodococcus jostii* RHA1 mit Hilfe solcher Sequenzen durchmustert und die interessanten Kandidaten identifiziert. *R. jostii* beherbergt eine ungewöhnlich große Anzahl oxidativer Enzyme, darunter mehr als 40 potentielle Flavoprotein Monooxygenasen. Der Fokus dieser Dissertation lag auf zwei Untergruppen der Klasse B Monooxygenasen: den Typ I Baeyer-Villiger Monooxygenasen (BVMOs) und den Flavin-beinhaltenden Monooxygenasen (FMOs). Zwar sind solche Enzyme bei Bakterien und Pilzen nicht selten, aber üblicherweise findet sich nur ca. ein BVMO Gen pro Mikroorganismus, FMOs sind sogar noch seltener. *R. jostii* stellt somit eine außergewöhnliche Quelle dar, und aufgrund der Kenntnisse über die vielfältigen Fähigkeiten dieses Organismus sind die Erwartungen an seine Enzyme hoch. Insgesamt wurden 23 BVMOs und sieben FMOs, sowie verschiedene andere Flavin abhängige Enzyme identifiziert, kloniert, in *E. coli* exprimiert (= produziert) und analysiert. In **Kapitel 3** werden die Ergebnisse der BVMO Analyse beschrieben und in **Kapitel 4 & 5** die der FMOs. Um die zeit- und bedingungsgleiche Analyse einer so großen Anzahl an Enzymen bewältigen zu können, wurden zwei neue Methoden entwickelt. Eine Methode dient der Bestimmung der Enyzmkonzentration NADPH abhängiger Enzyme im Zelllysat, also dem Gesamtinhalt einer Zelle nach deren Aufbrechen, was eine aufwendige und zeitraubende Aufreinigung der Enzyme überflüssig macht. Die zweite Methode stellt einen quantitativen Test zum Nachweis des NAD(P)H Verbrauchs dar, wodurch indirekt auf die Enzymaktivität geschlossen werden kann. Mit Hilfe dieses Tests kann ein Enzym gleichzeitig und schnell für sehr viele verschiedene Substrate getestet werden. Die Kombination beider Methoden erlaubt ohne großen Aufwand eine schnelle und dennoch

akkurate Analyse neuer Enzyme. Zusätzlich konnten wir aufgrund der großen Datenmenge ein neues Sequenz-Motiv für die Identifikation von BVMOs in Datenbanken identifizieren.

Die Resultat der FMO Analyse (**Kapitel 4**) ergab, dass es sich hierbei um eine eigene, neue Untergruppe der Klasse B Monooxygenasen handelt: Typ II FMOs. Alle sieben Kandidaten zeigen das typische FMO Sequenz-Motiv, katalysieren aber die Umsetzung typischer Substrate dieser Enzymklasse nur schlecht. Dafür sind sie – ganz im Gegenteil zu allen anderen Untergruppen der Klasse B Monooxygenasen – in der Lage, nicht nur NADPH, sondern auch das günstigere Coenzym NADH zu verwenden. Schon allein diese Fähigkeit macht sie zu äußert interessanten Biokatalysatoren, deren natürliches Substratspektrum dringend erforscht werden sollte. Drei Enzyme innerhalb dieser neuen Gruppe waren zusätzlich noch in der Lage typische Baeyer-Villiger Oxidationen durchzuführen, eine Reaktion, die von FMOs üblicherweise nicht ausgeführt wird – aber besonders zusammen mit der Coenzym Flexibilität eine äußerst attraktive Fähigkeit darstellt. Aufgrund dessen haben wir diese drei Kandidaten in **Kapitel 5** genauer untersucht, weitere Substrate getestet und ihre kinetischen Eigenschaften analysiert. Anhand eines Modells (siehe Figure 5-4) wurden die Aminosäuren des potentiellen aktiven Zentrums identifiziert und ihr Einfluss auf die Enzymaktivität durch Mutagenesestudien bestätigt. All diese Daten deuten auf eine höchst interessante neue Enzymklasse hin. Ein besonderes Merkmal scheint eine extralange Sequenz am Enzymende zu sein, mit deren Hilfe sich viele weitere potentielle Kandidaten in den Datenbanken finden lassen.

Neben der Entwicklung und Bereitstellung neuer Methoden zur Charakterisierung oxidativer Enyzme hat diese Dissertation nicht nur die Anzahl charakterisierter BVMOs drastisch erhöht, sondern stellt auch die Geburtsstunde eines neuen und vielversprechenden Mitglieds der Flavoprotein Monooxygenase Familie dar. Diese neuen Typ II FMOs haben das Potential die neuen Stars der Klasse B Monooxygenasen zu werden, sozusagen mit glänzend gelben Aussichten...

Bei Interesse steht hier eine drei-minütige Einführung zum Thema Enzyme zur Verfügung, live von meinem Auftritt bei “FameLab 2013” in Leipzig.

“Von Werkzeugkästen und Heinzelmännchen”

how to be an artist

stay loose. learn to watch snails. plant impossible gardens. invite someone dangerous to tea. make little signs that say yes! and post them all over your house. make friends with freedom and uncertainty. look forward to dreams. cry during movies. swing as high as you can on a swingset by moonlight. cultivate moods. refuse to be "responsible". do it for love. take lots of naps. give money away. do it now. the money will follow. believe in magic. laugh a lot. celebrate every gorgeous moment. take moonbaths. have wild imaginings, transformative dreams, and perfect calm. draw on the walls. read every day. imagine yourself enchanted. giggle with children. listen to old people. open up. dive in. be free. bless yourself. drive away fear. play with everything. entertain your inner child. you are innocent. build a fort with blankets. get wet. hug trees. write love letters...and dance as much as you can!

Unknown, often quoted as from Joseph Beuys or SARK – but true nevertheless

ACKNOWLEDGEMENTS

Finally done!

Done?

Probably not.

And even if it would be – I am sure I´d find something that could/should have been better in this or that way...I guess that´s life ☺

Something that could definitely not have been better are the people I have met in this wonderful city of Groningen!

But before I start getting lost mentioning all these great people, I want to thank some special ones: my PhD Supervisor Prof. Marco Fraaije for blindsiding me with the chance to make a PhD in the Netherlands. It was you, who inspired me to give science a new chance, to try this adventure of making a PhD abroad and who gave me the freedom to develop and find to myself within. You are a great supervisor – never stop being this friendly, easy going and young minded!

Actually, at that point I also have to thank Prof. Dick Janssen. Not only for having Marco around – and therefore also us "Fraaijos" – but especially for giving Sebastian the chance to spend this time together with me in Groningen. I think we both agree that my persistent reminders to deal with "this application from Germany" were a good idea…

However, all this could not have been finished without the members of the reading committee: Prof. Dr. L. Dijkhuizen, Prof. Dr. G.J.W. Euverink and Prof. Dr. W.J.H. van Berkel. Thank you for taking the time to go through these few (…) pages and the helpful suggestions. Also for taking part in my defence ceremony together with Prof. Dr. D.B. Janssen, Prof. Dr. J.G. Roelfes and Jun.-Prof. Dr. A. Schallmey. Here, I want to specially thank Anett. Du hast mich fast mein ganzes bisheriges wissenschaftliches Leben begleitet und warst mir immer eine große Hilfe! In Greifswald eher zufällig aufgrund der ähnlichen Namen einander zugeordent, setzte sich die Zusammenarbeit sogar bis in das "ferne"

Ausland fort – unabhängig voneinander geplant aber dafür umso schöner. Es fühlt sich einfach richtig an, dass du heute hier mit dabei bist. Vielen Dank!

From an organisational point of view, the next thank should probably go to my Paranymphs Ana and Gosia. But I do not only want them mentioned here because of their "Paranymph duty": dear Ana and Gosia, I am very grateful for our friendship!
"Paranimfen" are a nice Dutch tradition that support and escort the Promovendus during the defense ceremony. Nowadays it is a more symbolic and decorational function and less protective and scientifically relevant than it used to be. However, you should choose them carefully and often they represent your closest friends. Unfortunately the system allows only two of them… I am really glad that choosing this couple was a tough decision, meaning that I had more of such real friends that were beside me almost from the beginning! Like Hanna for example. I will not start making a list here, the respective people know… all I want to say to these special friends is THANK YOU for your company and for sharing your lives with me! We will continue and I am looking forward to that!

With finishing this book, a special thank goes to Geeske for the summary translation – without her, the "nederlandse leek" might have had a hard time reading selfsame book ☺.

Some other group also had a quite important influence on the outcome of this book – the 'stuff' behind the scenes: our technician Piet and the secretaries Sandra and Karlien. Only because of you everything can run so smoothly! We all know that, even though we probably do not often enough tell you so. That´s why I do it here: THANK YOU!

The next group I want to thank are my BIOMOX fellow ladies Evelien, Stephania, Harini – and the first lady of all: René. It was nice to meet you regularly throughout the year at different places in the Netherlands. And it was great to see all our projects develop and grow together to one big project. It is a pity that it ended when finally all our individual projects reached a stage where we eventually could work together, despite the distance

of our labs – but the way scientific research is organized it will always be like that: funding is over when the really interesting results finally come… It seems in this point there is a correlation between life and science… However, the end of one thing is just the beginning of something new and working with you was a nice experience to take along. I wish us all a happy continuation on whatever and wherever we want it!

Another few fellows that made these years special are my two and a half bachelor students Willem, Alrik and "half" Michael, and our apprentice Raoul. It was a pleasure to teach you and guide you through the lab – and to take part in your lives as well. You gave me the opportunity to work on a totally different aspect of science – the personal one. You showed me that I really like that. Thank you!

Next I want to thank all my office mates in the "baby office". You were not only a good address for scientific questions – here I especially want to thank Edwin and Danny – as the inofficial name of the office already implies, I also learned a lot about raising children in these few years ☺. It might well be that the surprisingly productive bird and tomato experiments are a result of the fertile environment – thank´s for tolerating my little side projects, and not letting them die during my holidays.

The group is of course much bigger than just one office and I really enjoyed the people we had around. Thank you for the nice atmosphere, the insight into your different cultures and sharing of your knowledge. Thank you for the many hours in the coffee corner or at restaurants, both, during lunch time at the campus or in the evenings in the city – or the best dinners of all, self cooked at home. Thank you for the borrels and private parties. Thank you for the excursions you took me along or joined my ideas. Thank you for all the support I received during the years and that you always agreed with my present suggestions (you cannot imagine how I felt when I heard "Anette´s pret-pakket" was a serious wish on an official wish list). Thank you for gladly having been my "culinary guinea pig" and always having liked my cakes ☺.

All these “Thanks” include of course also the numerous guests from all over the world we had in these years, as well as several members of the other groups in the “centre of ‘rondje lab’”!
Now substract the work related “Thanks” and apply the list to all the other friends I made here, mainly found in dutch classes, couch surfing events and “Expats of Groningen” or GRASP (the PhD student network, now called Gopher) activities.
I did not only spend my time with foreigners (even though that is often much easier to realize and organize than really meeting and getting to know a real Dutch…) I also spend some time with “natives”, mainly playing or dancing ☺. There is first to mention the huge balfolk community in the Netherlands that is really active throughout the whole country. Last but not least I want to mention the spellenvereiniging Rabenhaupt, where I spend most of my Sunday nights. To be honest, this is where I really learned my Dutch...

And then there is still the most important person of all: Lieber Sebastian, vielen Dank, dass du immer für mich da warst und bist, und zwar ganz egal wo. Viele Beziehungen überleben es nicht wenn mehrere Hundert Kilometer dazwischen liegen, bei uns hat es geklappt. Allerdings will ich auch nicht bestreiten, dass es sehr schön war, dass du mir letztlich ins Ausland gefolgt bist - deine Anwesenheit hat nicht unerheblich zu dem positiven Fazit über den Aufenthalt in den Niederlanden beigetragen ☺. Hier haben wir auch endlich das Experiment “zusammenwohnen” ausprobieren können – und ich glaube behaupten zu können, dass es ganz großartig funktioniert... Auch die Tatsache, dass dies jetzt schon das zweite wissenschaftliche Buch ist, in dem ich mich bei dir für deine Unterstützung und deine Geduld mit mir bedanken kann macht mir Mut für die Zukunft. Ich freue mich schon auf die nächsten gemeinsamen “Experimente”…

Auch meiner Familie möchte ich für ihre Unterstützung danken, sowie für die vielen Besuche, egal, in welchem Teil der Welt ich mich auch gerade aufhalte.

Finally I can just say, that this time in the Netherlands was great – and you all were part of this greatness. I will miss the canals in every city, everything that can be described with that unique dutch expression “gezellig”, the cows on every inch of grass, the cute language, the borrels, the old-fashioned windmills on the horizon, myriads of flowers, pepernoten, the market culture, the healthy working attitude, the gadget shops, the cozy villages,… and yes, even a kroket now and then. I learned so much besides science and language! Here I met the whole world in so many different ways and tastes.
Thank you just for having been at the same place where I was at the very same moment! For all the upcoming moments in your life I wish you all the best – and I hope we might share a few more ones together!
(You probably don´t know this yet, but I never forget anyone I like – that’s a promise. If you see that now as an honor or a challenge, well that’s up to you ☺.)

In the end, I would like to give you a small gift to remind you of me – as often as you want. I know how difficult the food situation in the Netherlands is, so just with you I will share a piece of my childhood: my very special recipe for my world famous, traditionally *pfälzer* (the palatinate, a special area in Germany), but personally perfectionized redwine cake that we shared on so many occasions. And as typically for any of my gifts it comes with lots of ingredients and comments that just together unfold its complete and complex (tasting) beauty….

That´s it finally (really), enjoy and think of me while eating...
):]*

Anette's world famous, traditionally pfälzer but personally perfectionized redwine cake

First grease a baking form and place it in the fridge (that cake loves to stick on the form, so this is just a try to trick it... at least I always had the impression that it works better this way - but maybe it's just superstition).

Mix 250g flour, 200g sugar, 2 tea spoons cinnamon (or probably more – as much as I know me... You can also add a hint of ground ginger, anise and ground gloves, I do that sometimes, depending on my mood on the baking day ☺), 2 table spoons cacao, 1 package baking powder and vanilla sugar each and 150g chocolate sprinkles (alternatively all sorts of crushed Nikoläuse / Sinterklaases / Santa Clauses or easter bunnys...).

In a separate bowl mix the wet stuff: 3 eggs, 250ml redwine and 1/3 of a 500g margarine package ☺ (you are old enough to calculate that yourself).
Now mix wet and dry stuff together. Cover the form (remember? its still in the fridge) with some flour and pour in the dough. Then bake it for around 1h at 180°C.

After its cooled down the original receipe now covers the cake with a layer of sugar icing (powdered sugar with a few spoons of redwine). For me the cake is already sweet enough, so I just sift a bit of powdered sugar over it to give it some homogenous colour. Another tasty alternative actually is a melted chocolate frosting.

And **NO**, I will not add any reference here. True enough, it is based on a traditional recipe from the german area "Pfalz" (where all the good german wines grows...), but the amounts of the sinlge ingredients are so much modified, that **I claim it to be my own invention!** (What makes this book now not only a thesis, but also an official reference for a cake - does that mean that its place is as much in the kitchen as it is in the lab?
Whatever, I think I like that idea, it's a great one ☺!)

Phantasie ist etwas, das sich manche Leute gar nicht vorstellen können.

Gabriel Laub

Curriculum Vitae

&

List of Publications

Anette Riebel was born on October, 15th 1983 in Worms 'am Rhein' (at the river Rhine), one of the oldest cities in Germany and scenery of the *Nibelungen Saga*. She grew up loving the historically and mythically impact all around what strongly influenced her predilections when choosing her myriads of hobbies. Since the age of four she was musically educated and proofed to be a talented pianist and singer. However, against all expectations she did not make music her profession but decided to study biochemistry.

She chose the Ernst-Moritz-Arndt University in Greifswald and moved to the other side of Germany after finishing the higher education (Eleonoren Gymnasium, Abitur) in 2002. Greifswald is placed close to the Baltic Sea and part of the former East Germany (DDR); now, fully restored, an absolutely adorable village. During the five years of studies she specialized on molecular biology (Prof. R. Walther) and biotechnology (Prof. U.T. Bornscheuer).

After receiving her diploma degree in 2007 she felt that it is time to also taste life outside the German borders and went to the Netherlands to continue her scientific career as a doctoral candidate. More precisely she went to Groningen, one of the most beautiful cities in the Netherlands. Within the BIOMOX project she explored different flavoenzymes in the Molecular Enzymology group of Prof. M.W. Fraaije at the Rijksuniversiteit Groningen. As the Netherlands are full of expats, these years where quite an international experience, leaving Anette with more impressions and worldwide connections than she ever thought of. She is very grateful for both, Dutch and international experience and not sure if she already had enough of it yet – the world is still so big and tempting…

At the moment she is back in Germany and working at the university of Leipzig – up to now the closest place to home since she left Worms 10 years ago. In her current position she is operating at the interface between research, industry and public by helping scientists to successfully start a business with their research results.

List of publications

D. E. Torres Pazmiño, **A. Riebel**, J. de Lange, F. Rudroff, M. D. Mihovilovic and M. W. Fraaije (2009) Efficient Biooxidations Catalyzed by a New Generation of Self-Sufficient Baeyer–Villiger Monooxygenases. *ChemBioChem* 10, 2595 - 2598.

A. Riebel, H. M. Dudek, G. de Gonzalo, P. Stepniak, L. Rychlewski, M. W. Fraaije (2012) Expanding the set of rhodococcal Baeyer–Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Applied Microbiology and Biotechnology* 95, 1479 - 1489.

A. Riebel, G. de Gonzalo, M. W. Fraaije (2013) Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. *Journal of Molecular Catalysis B, Enzymatic* 88, 20 – 25.

A. Riebel, M. J. Fink, M. D. Mihovilovic, M. W. Fraaije (2013) Type II flavin-containing monooxygenases: a new class of biocatalysts that harbors Baeyer-Villiger monooxygenases with relaxed coenzyme specificity. *ChemCatChem*, DOI: 10.1002/cctc.201300550

Scientific Awards

Biocat Poster Award at the Biocat 2008: International Congress on Biocatalysis (Hamburg, Germany).

The great tragedy of Science –
the slaying of a beautiful hypothesis by an ugly fact.

Thomas H. Huxley